Data Integration and the Accounting Labor Market Effects of Generative AI*

Chuck Downing
Sloan School of Management
Massachusetts Institute of Technology

June 2026

Abstract

This paper examines how generative AI (GAI) is reshaping demand for corporate accountants. GAI is a technological shock that requires complementary investments to realize productivity gains. I show that within-industry variation in prior investment in data integration—the extent to which information can flow across a firm's multiple internal systems automatically—predicts GAI adoption in the accounting function. I find that firms with one-standard-deviation-greater prior investment in data integration reduce accountant employment by 2.7% and job postings by 13% post-GAI. The declines are concentrated in junior positions, leaving demand for manager-level accountants stable, and are driven by attrition rather than layoffs. I then use the text of job postings and résumés to examine changes in the task content and skill requirements of remaining accounting jobs. I document a reduction in routine information processing tasks and a shift toward analytic and strategic tasks, along with wage premia for manager-level roles performing more expert work. Data-integrated firms increasingly seek candidates with CPAs, technical backgrounds, and advanced degrees. Taken together, the evidence is consistent with GAI reducing the demand for junior accountants and pushing the profession toward higher-skill roles requiring greater accounting expertise.

* Email: downingc@mit.edu. I am grateful to my dissertation committee, Eric So (co-chair), Andrew Sutherland, Rodrigo Verdi (co-chair), and Joe Weber for their invaluable guidance. I thank Darren Bernard, Beth Blankespoor, John Core, Anna Costello, Ed deHaan, Joao Granja, Michelle Hanlon, Steve Kachelmeier, David Kim, S.P. Kothari, Anthony Le, Becky Lester, Joe Moran, Chris Noe, Jim Omartian, Matt Phillips, Jeffrey Saviano, Nemit Shroff, Kevin Smith, Chris Stewart, Mitchell Troyanovsky, Felix Vetter, Charles Wang, Chloe Xie, and employees at IBM, EY, and SAP for providing helpful comments and suggestions. I also thank seminar participants at Chicago Booth, Harvard Business School, Miami, Michigan, MIT Accounting, the MIT BIG.AI conference, the MIT economics of AI reading group, Rice, Stanford, UT Austin, Washington, as well as participants at the 2025 WashU Olin Accounting Research Conference. I am grateful to MIT and the Deloitte Foundation for their generous financial support. All errors are my own.

1. Introduction

This paper examines the labor market consequences of generative AI (GAI) in corporate accounting. GAI can improve individual productivity in many settings (Noy and Zhang, 2023; Brynjolfsson et al., 2025b; Choi and Xie, 2026), but large-sample empirical work has not found clear changes in labor demand attributable to GAI (Humlum and Vestergaard, 2025a; Frank et al., 2026). This pattern is characteristic of the early stages of technological change, where productivity gains depend on complementary investments in processes and infrastructure (Solow, 1987; Bresnahan and Trajtenberg, 1995). As such, the labor market effects of GAI are unlikely to be uniform across firms because adoption at scale requires these organizational complements. This study identifies *data integration*—the extent to which information can flow automatically across a firm's internal systems—as a critical complement for GAI in accounting. I measure firms' prior investment in integration relative to industry peers and test whether the demand for corporate accountants shifts after the emergence of GAI at firms with greater pre-existing capacity to adopt it.

For GAI to perform tasks typically handled by humans, it must access the same organizational knowledge within existing workflows (Davenport and Tiwari, 2024; Salesforce, 2025). Data integration makes this knowledge accessible and actionable (Bughin et al., 2023). Integration is one dimension of a firm's broader internal information environment (Gallemore and Labro, 2015). It is especially consequential for GAI because the technology can only act on the information a firm's systems can feed it. Integration is particularly relevant in corporate accounting, where work requires processing information from disparate systems—subledgers, bank statements, inventory, payroll, tax, CRM—with the general ledger.[1] GAI can draft reconciliations, conduct analyses, and create documentation, but only data-integrated firms

[1] Enterprise Resource Planning (ERP) systems are meant to connect these systems. However, most organizations still face fragmented data flows and siloed systems even within their ERP, meaning that GAI models often cannot access the required enterprise-wide information systematically (Salesforce, 2025).

can supply the context to deploy these capabilities in enterprise workflows. By contrast, fragmented firms are limited to narrow, voluntary chatbot use by individual employees (Valentine et al., 2025; Bloom et al., 2014). Prior integration investments thus provide ex ante variation in adoption capability that I exploit to study the labor market consequences of GAI in corporate accounting.

To measure investment in data integration, I use résumé data from Revelio Labs to calculate the employment of data management professionals as a share of total white-collar employment at a given firm.[2] These employees are responsible for the integration and centralization of organizational data across business units. A greater prevalence of these roles signals sustained investment in making internal data accessible and usable—the foundation GAI requires to operate across a firm's systems (Davenport and Tiwari, 2024). The approach builds on labor-based measures of firms' data assets (Abis and Veldkamp, 2024; Eisfeldt et al., 2023), but is meant to capture data accessibility rather than quantity—whether internal data can flow across systems, not how much of it exists.

I study the impact of GAI on corporate accountant demand through the lens of the task framework of Acemoglu and Autor (2011). Firms allocate tasks between labor and capital; new technologies lower the quality-adjusted price of capital for performing specific tasks (Kogan et al., 2023). Labor demand then adjusts through three channels: (i) displacement, as capital directly substitutes for tasks previously done by workers; (ii) productivity, as workers perform some remaining tasks more efficiently; and (iii) reinstatement, as the technology creates new tasks. The net effect also depends on the elasticity of demand for the occupation's output: productivity gains raise employment only if output expands in response (Bessen, 2019). Whether GAI reduces or increases demand for accountants depends on the magnitude of these three channels and the nature of the tasks that remain for labor (Acemoglu and Restrepo, 2018; Autor and Thompson, 2025).

[2] O*NET codes 15-1242 (Database Administrators) and 15-1243 (Database Architects).

Several features of corporate accounting shape how these channels operate. Accounting is a cost center whose output is largely fixed by reporting and compliance requirements; because public firms may not demand more accounting when it becomes cheaper to produce, productivity gains are likely to reduce headcount rather than expand output. At the same time, low error tolerance necessitates human oversight, tempering the pace of near-term displacement. And because accounting draws on internal data from across the firm, data integration is a relevant complement to GAI adoption relative to other occupations. While these institutional features motivate both my focus on data integration and my empirical approach, the timing, magnitude, and distribution of effects remain empirical questions.

I examine accountant labor demand before and after the emergence of GAI, marked by ChatGPT's release in November 2022. My sample consists of 3,620 publicly traded firms from 2010–2025, constructed by linking résumé and job posting data from Revelio Labs to firm financial data from Compustat. I measure a firm's pre-existing capacity for GAI adoption using its 2022 industry-adjusted employment share of data management professionals, fixed before firms could respond to GAI. A natural concern is that this measure captures general technological sophistication or managerial quality rather than data integration specifically, and that these traits correlate with pre-existing trends in accountant demand. To address this, I control throughout for the firm's 2022 industry-adjusted share of non-data IT workers; in robustness tests, I also use this IT share as a placebo treatment. My strategy thus uses firm-level variation in prior integration investment to capture differential exposure to a new general-purpose technology.

I first validate that data-integrated firms exhibit greater GAI adoption by examining post-ChatGPT mentions of GAI across four sources. A one-standard-deviation increase in data integration is associated with a 4% increase in 10-K mentions, a 22% increase in earnings call mentions, a 16% increase in job posting mentions across all occupations, and an 11% increase in employee résumé mentions. Each source is subject to different reporting incentives; 10-Ks and earnings calls face AI-washing concerns (Barrios et al., 2025), while job postings

and résumés reflect hiring intentions and employees' self-disclosed work. The agreement across sources supports the interpretation that data integration creates meaningful differences in GAI adoption rather than differences in disclosure behavior. Disaggregating by occupation, the pattern holds within the accounting function specifically: a one-standard-deviation increase in data integration corresponds to an 11% increase in GAI mentions in accounting job postings and a 10% increase in accountant résumés.

My main analyses show that accountant employment declines at data-integrated firms following ChatGPT's release. I study both the "core" accounting occupation (accountants and auditors) and the broader accounting and finance function.[3] A one-standard-deviation increase in data integration is associated with a 2.7% reduction in core staff-level accounting employment (2.3% for the broader function), with effects strengthening over time to 4.7% by 2025. To control for firm-wide labor demand shifts, I compare employment changes within the same firm-year to legal employees using a triple-difference design. Lawyers are similarly exposed to GAI (Handa et al., 2025) and are a common group for comparison to accountants (e.g., Bloomfield et al., 2017), but legal GAI tools depend less on data integration across business units. This within-firm comparison yields a significant relative decline in accounting versus legal employment, confirming that the decline is specific to accounting rather than reflecting firm-level shocks coincident with ChatGPT's release.

I supplement the résumé-based analysis with job posting data to capture forward-looking changes in labor demand. A one-standard-deviation increase in data integration is associated with a 13% reduction in core accounting job postings and an 8% reduction in broader accounting postings. These magnitudes exceed the employment effects, consistent with postings capturing intended future hiring reductions before they materialize in headcount (Davis et al., 2013). Decomposing the employment decline into hiring and separations reveals

[3]Bookkeeping and receivables clerks, financial analysts, financial managers, risk specialists, and other financial employees.

the same pattern: the reduction is driven entirely by a sharp decline in new hires (8–8.5%), with no significant change in separations. Data-integrated firms reduce accounting headcount by curbing replacement hiring rather than through layoffs. Together, these margins indicate that labor demand adjusts ahead of fully realized productivity gains, consistent with firms responding to anticipated automation.

Next, I document a compositional shift within data-integrated firms toward senior accountants using a triple-difference design comparing managers and staff at the same firm in the same year. A one-standard-deviation increase in data integration is associated with a 3.6% increase in manager employment relative to staff, strengthening to 5.7% by 2025. This differential is driven by staff declines rather than managerial growth—manager headcount is stable in the baseline specification. Examining disclosed wages in job postings, I find that salaries for manager-level postings at data-integrated firms are 4–6% higher per standard deviation of data integration, while staff-level wages are unchanged. Together, the employment and wage patterns are consistent with GAI displacing junior workers while augmenting the value of senior expertise (Autor and Thompson, 2025), and with growing evidence that GAI's labor market effects concentrate among junior workers (Brynjolfsson et al., 2025a; Lichtinger and Hosseini Maasoum, 2025).

My empirical design faces three main identification concerns. First, firms select into data integration investment, and the firm characteristics driving that selection may independently predict changes in accounting employment. Second, data integration may reduce accounting headcount mechanically and permanently—integrated systems may simply require fewer accountants—so that the post-2022 declines reflect a general consequence of integration rather than a GAI-specific channel. Third, employment of data management professionals may reflect data *problems* rather than data capabilities: firms might hire integration specialists because their systems are fragmented, and GAI may then reduce accounting labor demand by helping these firms resolve their integration problems rather than by automating accounting work at well-positioned adopters.

Several pieces of evidence help address these concerns. To address selection, I re-weight the sample using entropy balancing on observable covariates, with inferences unchanged. To examine a permanent main effect of integration, I replace the 2022 treatment with a 2019 base year: if integration always reduced accounting headcount, the earlier measure should predict employment declines before ChatGPT's release. It does not—event-time estimates are flat through 2022—and placebo treatments centered on 2016 and a rolling average of investment likewise produce no effects. Two further patterns weigh against the data-problems interpretation. In determinants tests, data-integrated firms exhibit higher-quality information environments, and the integration measure is sticky over time—consistent with a stable firm type rather than remedial spending on broken systems. The validation results show greater GAI adoption at these firms across disclosures, job postings, and résumés, concentrated within the accounting function—in line with firms deploying the technology rather than being rescued by it. Finally, an alternative treatment based on résumé mentions of specific integration software yields directionally consistent effects.

Having documented a decline in demand for junior accountants, I examine changes in the task content of remaining accounting jobs, using a large language model to classify the text of individual job postings.[4] Remaining accounting work shifts toward more complex tasks: routine task intensity falls significantly, while the ratios of analytic-to-routine and strategic-to-routine tasks rise significantly. These shifts indicate that data-integrated firms are recomposing accounting jobs away from the routine work GAI substitutes for and toward analysis and judgment, especially within the junior roles that remain.

These changes suggest that GAI-related automation raises the expertise required of remaining accountants (Autor and Thompson, 2025). Consistent with this, staff-level job

[4] For each core accounting job posting, I instruct the LLM to generate a binary indicator for each of the 29 O*NET task statements listed for Accountants and Auditors (O*NET 13-2011). I group these binaries into routine, analytic, and strategic categories; tasks not clearly belonging to one of these form a residual used only as a control. Details of the groupings are in the online appendix.

postings at data-integrated firms become more likely to require or prefer a CPA, while manager-level postings show no change. I then analyze the résumés of the accountants entering these firms and show that new accounting hires are significantly more likely to hold CPA licenses, accounting degrees, technical degrees, and master's degrees, consistent with a rising skill threshold for entry into the profession and with demand for hybrid skill sets combining accounting knowledge with technical capacity.

Taken together, my results are consistent with a skill-biased reorganization of corporate accounting work. The first contribution of this paper is to show why the labor market consequences of GAI differ so much across firms: while GAI raises productivity on accounting tasks in controlled settings (Choi and Xie, 2026), large-sample work finds little change in overall employment (Humlum and Vestergaard, 2025a)—a gap that makes sense if effects concentrate only at firms with the organizational complements to start adopting GAI at scale. I show that data integration, a distinct component of the internal information environment, is a critical complementary investment for moving from individual AI experimentation to enterprise-level adoption, consistent with the beliefs of practitioners (Valentine et al., 2025; Salesforce, 2025). The quality of a firm's internal information environment thus shapes the incidence of technological adoption, with implications for how organizations adjust to and create value from technological disruption.

In addition, I provide new evidence on the distributional effects and changing skill requirements of accounting jobs in response to GAI. Staff-level accountants are particularly affected, consistent with GAI substituting for the work of lower-rank white-collar employees (Berger et al., 2024; Brynjolfsson et al., 2025a; Lichtinger and Hosseini Maasoum, 2025). I also provide direct evidence that corporate accounting jobs are shifting toward analytic and strategic work. While concurrent work links investments in predictive AI to reduced demand for accounting skills (Böke et al., 2024), I show that GAI adoption at data-integrated firms reallocates remaining work toward tasks that *increase* the value of accounting expertise, credentials, and hybrid skill sets combining accounting knowledge with technical

capacity—consistent with theories of automation that emphasize rising expertise requirements in the tasks that remain for humans (Autor and Thompson, 2025). These results speak to what forms of human capital are most complementary to GAI in accounting (Friedman et al., 2024; Fedyk et al., 2022; Law and Shen, 2024), informing educators and policymakers on the future of accounting work.

2. Institutional Background

Accountants draw on data from across the organization—sales and procurement systems, bank statements, payroll and human resources records, production and inventory logs, and capital expenditure tracking—to process information into a coherent financial record. This work relies on a layered technology stack: ERP systems serve as the central ledger, specialized modules support functions like procurement and payroll, and business intelligence tools support aggregation, reconciliation, and analysis. These systems constitute a core component of the firm's internal information environment, and prior research links their quality to outcomes including financial reporting quality, internal control effectiveness, and the timeliness of reporting and decision-making (Gallemore and Labro, 2015; Morris, 2011; Dorantes et al., 2013; Charoenwong et al., 2024; Heese and Pacelli, 2025; Ashraf, 2025).

Although these systems provide a common infrastructure, accounting work still involves substantial manual effort (Andrusko and Amble, 2024). Accountants reconcile records across modules, make adjusting entries, and compile documentation and narrative disclosures. These tasks persist because accounting requires judgment, interpretation, and assurance—areas where rules-based process automation and traditional predictive AI (e.g., machine learning) offer limited support. Traditional AI typically handles narrow tasks that take structured data as input, such as anomaly detection, fraud monitoring, or forecasting. Prior research examines how traditional AI affects accounting labor markets in both audit (Fedyk et al., 2022; Law and Shen, 2024) and corporate accounting (Böke et al., 2024).

Generative AI extends the accounting toolkit by supporting work beyond structured, rule-based tasks. Unlike predictive systems relying on tabular and structured data sources, large language models can produce contextually relevant output that previously required human reasoning. Their strength lies in contextual interpretation—summarizing unstructured documents, reconciling discrepancies across sources, and drafting narrative explanations that reflect underlying data. GAI supports the judgment-oriented aspects of accounting, representing a step-change from prior automation. This makes it especially useful for reconciliation, information retrieval, and preparing documentation and disclosures—tasks that form a substantial portion of human accounting work.

Early organizational use of GAI typically involves stand-alone chatbots and co-pilots that provide natural language interfaces for queries or simple drafting. These tools lower search costs and produce quick summaries, but they operate outside of core systems and serve as individual productivity aids rather than structural changes to how accounting work is performed. Because usage is determined ad hoc by individual employees, agency problems may prevent firms from capturing the productivity benefits of these tools (Humlum and Vestergaard, 2025a). Full integration requires embedding GAI directly into the accounting workflow. The most valuable applications—automated reconciliation, documentation, and analysis—depend on access to the same information human accountants use, requiring systematic exposure to data across ERP systems, subledgers, and related business systems. Unlike human employees who can navigate systems piecemeal, generative models need structured access through secure data pipelines, demanding new infrastructure for integration, validation, and control (Valentine et al., 2025). With these foundations, GAI can move beyond surface-level support to become an integral component of the firm's information environment.[5]

[5] Firms adopt GAI in accounting through three main channels: (i) GAI tools from ERP providers such as SAP, NetSuite, or Dynamics 365; (ii) third-party ERP plug-ins targeting specific tasks; or (iii) in-house models built on proprietary data (least common). Adoption often begins with pilots in reconciliation or reporting before scaling across workflows. Each approach balances cost, customization, and control, but effectiveness ultimately depends on the quality and accessibility of underlying data regardless of choice.

Enterprise GAI adoption depends on foundational data infrastructure (Bughin et al., 2023; Davenport and Tiwari, 2024). Organizations that invested in data quality and architecture to support information flows and predictive analytics are better positioned to adapt these systems for GAI. However, integrating high-quality data across complex systems, orchestrating data flows, and aligning outputs with business needs remains a substantial challenge (Valentine et al., 2025). Firms with stronger pre-existing data integration face lower costs when scaling enterprise adoption beyond the pilot phase where many organizations remain (KPMG, 2024; Challapally et al., 2025). Finally, despite its promise, GAI in accounting still requires strong governance, validation, and human oversight. Automated systems can accelerate data collection, reconciliation, and documentation, but assurance still requires human judgment, especially when information is incomplete, conflicting, or sensitive. GAI is best understood as supporting preliminary information gathering and processing, leaving final interpretation and decision-making to human accountants—a division of labor that reflects both regulatory requirements and the continuing importance of professional judgment in financial reporting.

3. Hypothesis Development

3.1. Early Adoption of General-Purpose Technologies

Generative AI is widely considered a general-purpose technology (Eloundou et al., 2024). Despite documented productivity effects in many settings (e.g., Noy and Zhang, 2023; Brynjolfsson et al., 2025b; Choi and Xie, 2026) and unprecedented corporate investment (Bick et al., 2026; Humlum and Vestergaard, 2025b), aggregate labor effects attributable to GAI have been challenging to identify (Humlum and Vestergaard, 2025a; Brynjolfsson et al., 2025a; Frank et al., 2026). This pattern is consistent with prior general-purpose technologies, whose productivity gains depend on complementary investments in processes and infrastructure that diffuse slowly and unevenly across firms (Greenwood et al., 1997; Helpman and Trajtenberg, 1994). Because GAI models are broadly accessible but their returns depend on firm-level

integration into existing workflows (McAfee et al., 2023; Lichtinger and Hosseini Maasoum, 2025), early effects should concentrate among firms whose pre-existing capabilities lower the cost of adoption (Bresnahan et al., 2002; Schubert, 2025).

I examine one specific complementary investment in the context of GAI that is relevant for accounting: data integration (Davenport and Tiwari, 2024). GAI models excel at synthesizing unstructured information and generating contextually relevant outputs—but their effectiveness in corporate applications depends on access to relevant organizational knowledge. Internal data integration—the systems, processes, and architectures that govern how firms collect, store, integrate, and retrieve their operational data—is what makes this organizational knowledge accessible and actionable for GAI systems. Firms with poor data integration face higher information retrieval costs (Bloom et al., 2014), reducing the return on investment in GAI because they cannot capture as much of the gains in information processing created by the new technology. Specifically, firms with a greater degree of data integration will find it less costly to move from ad hoc experimentation by individual employees to enterprise-level and scalable adoption embedded in existing workflows (Valentine et al., 2025). Differences in outcomes between these "data-integrated" and "data-constrained" firms offer a way to partially isolate the effects of GAI during its early diffusion, forming the basis of this study's approach.

3.2. The Labor Market Consequences of Technological Change

I use pre-existing differences in data integration to examine how GAI affects the demand for accounting labor at public companies. Following the task framework (Autor et al., 2003; Acemoglu and Autor, 2011), I conceptualize jobs as bundles of tasks performed by labor or capital, where technological innovation reduces the cost of capital for specific tasks (Kogan et al., 2023). I assume that firms with greater prior investment in data integration experience a larger technological shock, as GAI embedded in integrated systems can perform or support more accounting tasks than GAI accessed as a standalone chatbot or co-pilot.

Labor demand adjusts in response to a technological shock through three channels (Acemoglu and Restrepo, 2018). Displacement occurs when GAI automates tasks previously performed by accountants, directly substituting capital for labor. A productivity effect emerges when GAI enhances accountants' efficiency in remaining tasks such as drafting reports or supporting decisions. Reinstatement occurs when GAI creates *new* tasks requiring human expertise, such as AI oversight. These channels operate through quantities (employment and new postings) and prices (wages): displacement exerts downward pressure, while productivity and reinstatement raise the marginal product of AI-complementary skills (Autor and Thompson, 2025; Hampole et al., 2025). The net effect will depend on the relative magnitudes of the three channels and on the elasticity of demand for accountant output.

Several features specific to corporate accounting shape how these channels may operate. Corporate accounting functions primarily as a cost center focused on compliance and accuracy rather than revenue generation, so productivity gains are more likely to translate into lower headcount than into expanded demand. Its low tolerance for error also limits experimentation and places a premium on reliability, which may slow near-term displacement. In addition, because accounting depends on operational and financial data drawn from across the firm's business units, data integration is an especially important complement to GAI adoption relative to many other occupations. Taken together, these features suggest that data-integrated firms will reduce overall demand for corporate accounting labor after the introduction of GAI, with the effects likely to appear more strongly in forward-looking hiring decisions than in the retention of existing employees.

Hypothesis 1: *Data-integrated firms reduce demand for corporate accountants relative to less integrated firms following the introduction of GAI.*

This reduction in labor demand is unlikely to be uniform across the organizational chart (Lodefalk et al., 2026). The tasks most susceptible to automation by GAI are routine information processing tasks that form the core responsibilities of junior accountants. In

contrast, senior roles such as controllers and accounting managers involve a greater share of tasks that require judgment, oversight, and interpersonal skills. While GAI may still serve as a productivity tool, it is less likely to substitute for these workers' core functions, reducing the potential for displacement (Demirer et al., 2026; Gans and Goldfarb, 2026). Therefore, I anticipate that the displacement effects of GAI will be concentrated in junior roles.

Hypothesis 2: *After GAI, data-integrated firms shift their labor demand away from junior and toward senior accountants, relative to less integrated firms.*

As GAI commodifies routine information processing, the role of accountants will transform even within a given seniority level. If tasks requiring less expertise are displaced, the average complexity of remaining work rises, shifting the focus from transactional duties toward strategic analysis, critical judgment, and decision-making (Autor and Thompson, 2025; Friedman et al., 2024; Kim et al., 2025). This shift will most significantly affect junior and mid-level professionals, who perform a higher proportion of routine work (Berger et al., 2024).

Hypothesis 3: *Within seniority levels, data-integrated firms shift remaining accountants away from routine tasks and toward strategic and analytical work.*

A change in the task content of jobs implies a corresponding shift in the skills firms value (Deming and Silliman, 2025). Concurrent work on predictive AI finds reduced demand for accounting-specific skills (Böke et al., 2024), but this may reflect the displacement of routine tasks typically associated with accounting work rather than a true reduction in demand for accounting expertise. If remaining tasks involve interpreting complex transactions, ensuring compliance with evolving standards, and exercising judgment in ambiguous contexts, demand for domain expertise may actually increase (Choi and Xie, 2026). Alternatively, demand may shift toward data analysis, software proficiency (Friedman et al., 2024), or soft skills like communication and teamwork (Ham et al., 2025), depending on which capabilities best complement the surviving bundle of work.

The wage implications depend on how automation reshapes the expertise threshold of

remaining work. Occupational licensing and credentialing already segment the accounting labor market by restricting entry into higher-value tasks. If automation removes routine functions, the average expertise required rises, increasing the scarcity and value of licensed professionals and raising wage premia (Autor and Thompson, 2025). If instead automation displaces more complex functions, or the surviving tasks become less accounting-specific and more general, the expertise barrier weakens, broadening the pool of qualified workers and compressing wages (Friedman et al., 2024). The wage effects of GAI thus depend not on the quantity of tasks automated but on whether the surviving task bundle raises or lowers the expertise required to participate in the occupation (Autor and Thompson, 2025).

Hypothesis 4: *Data-integrated firms increase the expertise requirements of remaining accounting positions.*

4. Data and Empirical Design

4.1. Sample Construction

I construct a firm-year panel from 2010–2025 merging résumé and job posting data from Revelio Labs with financial data from Compustat, returns from CRSP, disclosure data from the WRDS SEC analytics suite, and guidance from IBES. I extend the sample back to 2010 to provide a long pre-GAI period for modeling temporal trends, as in Blankespoor et al. (2026) and Berger et al. (2024), but not earlier due to reduced résumé coverage. Revelio Labs provides granular workforce data primarily sourced from LinkedIn profiles. I process position-level data for U.S. employees at public firms with valid start/end dates, creating individual-year observations. Manager-level positions are defined as seniority level > 3 (manager to senior executive, per Revelio's classification)[6], and I require individuals to be at a firm for six months to be counted in that firm-year, as in Hampole et al. (2025).

[6] Revelio's ensemble seniority metric combines current and past job characteristics, employment history, and age to predict the seniority of a given role.

I aggregate the data to the firm-year level using Revelio's adjustment weights to count accounting employees. I measure both the core accounting function (O*NET code 13-2011: Accountants and Auditors) and the broader accounting and finance function.[7] For within-firm comparisons, I restrict to white-collar occupations based on leading two-digit O*NET codes.[8] Missing firm-year observations are filled with zeros to create a complete panel capturing employment dynamics, including non-employment.[9] I also use Revelio's COSMOS job posting data, sourced from company websites, major job boards, and staffing firm sites. The data is de-duplicated and parsed into standardized fields including company identifiers, posting and removal dates, estimated salary, seniority, job taxonomies, location, and full post text. Meaningful job posting coverage begins in 2016, so firm-year posting-count analyses span 2016–2025; analyses that rely on the full text of individual postings (task content, CPA requirements, and posted wages) are restricted to 2020–2024, where text coverage is densest.

4.2. Measuring Data Integration

I measure data integration using a labor-based approach similar to Abis and Veldkamp (2024) and Eisfeldt et al. (2023). My main measure captures salary-weighted employment shares of database administrators (O*NET 15-1242) and database architects (O*NET 15-1243), demeaned within an industry (3-digit NAICS):

$$\text{DataIntegration}_{i,2022} = \frac{\sum \text{DataIntegratorSalaries}_{i,2022}}{\sum \text{TotalWCSalaries}_{i,2022}} - \underbrace{\frac{\sum_{j\in\mathcal{I}} \text{DataIntegratorSalaries}_{j,2022}}{\sum_{j\in\mathcal{I}} \text{TotalWCSalaries}_{j,2022}}}_{\text{Industry Mean (3-digit NAICS)}}$$

[7] SOC codes: Financial Managers (11-3031), Financial/Investment Analysts (13-2051), Financial Risk Specialists (13-2054), Actuaries (15-2011), Billing/Posting Clerks (43-3021), and Bookkeeping/Accounting/Auditing Clerks (43-3031). These roles largely handle internal financial information.

[8] I keep O*NET codes starting with 11-29 (Management, Business, Science, and Arts Occupations) and 41-43 (Sales and Office Occupations), consistent with Eisfeldt et al. (2023).

[9] A concern underlying this approach is that Revelio's data availability may correlate with underlying firm characteristics. Zeroes are imputed for roughly 10% of accounting observations, and about 20% of data integration observations. All main results remain robust to requiring non-missing observations, and/or a balanced panel of non-zero-only firms.

I use predicted salaries rather than employee counts because salary better captures employee productivity and value and is unlikely to be biased in the cross-section of data integration employees. Scaling by total white-collar salaries creates an intensity measure independent of firm size. Because industries differ in IT and data management requirements (Charoenwong et al., 2024; Abis and Veldkamp, 2024), I demean by industry to capture relative firm readiness for GAI. I choose NAICS because it is a standard product market-based taxonomy that allows for well-populated groups at the 3-digit level, but obtain economically similar results with alternative classifications, as well as without demeaning.

I fix the main measure at its 2022 value (pre-ChatGPT) to mitigate endogeneity, ensuring treatment reflects pre-existing capabilities rather than responses to GAI. Using 2022 employment rather than a rolling average is appropriate because GAI's benefits require human capital capable of adapting data flows to support unstructured data and retrieval-augmented generation (Valentine et al., 2025), rather than just investment in pre-existing architecture.[10]

My sample consists of 3,620 publicly listed U.S. firms linked to Revelio. Table 1 reports descriptive statistics. The average firm pays just under 1% of their total white-collar salaries to employees responsible for data integration across business functions. My main measure has a mean of zero by design, as it is de-meaned by industry. A one-standard-deviation increase in data integration corresponds to total white-collar salaries paid to database administrators and architects increasing by 0.8 percentage points within a given firm, representing substantial variation in investment in data integration. Figure 1 plots the cross-sectional distribution of the measure, which is approximately symmetric around its industry mean with meaningful mass in both tails.

Panel B of Table 1 reports cross-sectional regressions of the 2022 measure on firm characteristics. Firms that invest more in data integration are larger, operate across more

[10]Robustness analyses confirm that the results are not sensitive to using a rolling average, or to fixing at 2019 investment levels or earlier.

business segments, carry lower book-to-market ratios, and file their 10-Ks more promptly, with most of these relationships holding both with and without industry fixed effects. Across the broader cross-section they also tend to be younger and to issue management guidance more accurately, though these associations attenuate within industry. Notably, data integration does not line up with higher profitability—return on assets enters negatively—indicating that the measure reflects investment in information infrastructure rather than general firm performance. I control for these characteristics throughout the analysis.

4.3. Empirical Design

To examine GAI's impact based on pre-existing investment in data integration, I employ a continuous difference-in-differences (DiD) design with an exposure intensity treatment (Callaway et al., 2024), defined as:

$$Y_{it} = \beta(\text{PostGPT}_t \times \text{DataIntegration}_{i,2022}) + \gamma \mathbf{X}_{it} + \delta(\text{Post}_t \times \mathbf{X}_{it}) + \alpha_i + \lambda_t + \varepsilon_{it}$$

where Y_{it} represents accounting employment outcomes for firm i in year t. PostGPT_t is an indicator for 2023–2025 and $\text{DataIntegration}_{i,2022}$ is the measure of data integration intensity discussed above. $\mathbf{X}_{it}$ is the vector of controls, and α_i and λ_t are firm and year fixed effects. I cluster standard errors mainly at the firm level. Firm-level controls include total assets, scaled capital expenditures, sales growth, firm age, leverage, profitability (ROA), average returns, and book-to-market ratio. I also control for accounting employment share (core accounting as a percent of white-collar employment) to ensure results do not reflect accounting department inefficiency. Appendix A provides detailed variable definitions. I also include IT employee intensity as a control to isolate data management capabilities from general technological sophistication.[11] This cross-sectional approach may also reduce concerns that factors in 2022

[11] IT occupations include: Computer Systems Analysts, Information Security Analysts, Computer Network Support Specialists, Computer User Support Specialists, Computer Network Architects, and Network and Computer Systems Administrators. Results are robust to excluding this control or using software employees

other than generative AI affected certain occupations (Frank et al., 2026).

I estimate models with count-like dependent variables (number of employees, job postings) using Poisson pseudo-maximum likelihood (PPML), which better handles zeroes (Cohn et al., 2022). For other outcomes, I use ordinary least squares (OLS). Causal interpretation relies mainly on a parallel trends assumption typical of DiD designs using cross-sectional exposure as an instrument. I assume that, absent GAI, firms with different data integration investments would exhibit similar trends in accounting employment, conditional on controls and fixed effects. In addition, my design requires that data integration is pre-determined with respect to the introduction of ChatGPT, and that it affects post-period accounting employment only by scaling firms' responsiveness to the GAI shock, rather than through correlations with unobserved variables that also affect accounting employment in the presence of GAI.

There are two specific concerns that threaten my identification approach. First, the parallel trends assumption could fail if investment in data integration systematically predicts future accounting employment changes regardless of GAI. For example, it may be correlated with initiatives to automate or offshore accounting work, or general trends to reduce headcount at a given firm. I address this concern through pre-trends analysis and placebo tests using pre-period data. If investment in data integration in period t always predicted employment reductions in future periods, these patterns would be observed prior to the introduction of GAI. Figure 2 confirms statistically parallel pre-trends, and pre-period placebo tests in Table OA.1 reveal no significant systematic relationship between data integration investment and subsequent reductions in accounting headcount.

Second, data integration might correlate with other firm characteristics that either enable GAI adoption in accounting or shape responses to macroeconomic events around 2022, such as managerial quality, technological sophistication, or propensity to invest in new technologies. I take two approaches to alleviate this concern. First, I control for non-data IT employee

instead. The correlation between Data Integration and IT intensity is 0.4.

intensity (constructed identically to the treatment variable), which should capture general technological sophistication separate from data integration. This control shows no consistent independent effect (online appendix) and does not attenuate my main results. I address concerns that changes in the market for accountants occurring around 2022 created differential incentives for data-integrated firms to automate accountants through channels other than their exposure to GAI. Choi et al. (2024) document no significant changes in real wages for accountants around 2022. This result, combined with my empirical design choices, reduce concerns that trends in the demand for accountants may have been unstable in the post period between firms with different degrees of integration, if not for GAI.

In addition, I use lawyers as a placebo occupation. Lawyers are comparable to accountants as credentialed, white-collar professionals (Bloomfield et al., 2017; Cascino et al., 2021) with similar GAI exposure (Eloundou et al., 2024; Handa et al., 2025). However, legal AI tools use static documents and do not require systematic data integration across enterprise systems, because they can operate in silos. If my results reflected these general firm characteristics rather than data integration specifically, then my instrument should reveal similar effects for lawyers. I find no significant impact on legal AI adoption or employment within data-integrated firms relative to other firms, supporting the interpretation that my results operate through the data integration channel.

4.4. Equilibrium

Finally, I note that this analysis exploits cross-sectional variation in a short window around the introduction of GAI. As such, the estimates are best interpreted as partial equilibrium effects operating primarily on the demand side. There is a large body of literature suggesting that labor demand responds more quickly to technological change than labor supply does (e.g., Katz and Murphy, 1992; Autor et al., 1998; Card and Lemieux, 2001). Firms can restructure hiring, adjust tasks, and reallocate headcount relatively quickly in response to a productivity-enhancing technology, whereas human-capital supply—through training,

credentialing, and occupational switching—adjusts more slowly to the new skill requirements that technology creates (Freund and Mann, 2026). This asymmetry in adjustment speeds means that short-run labor market effects of technological change manifest primarily as wage and employment differentials across worker types, with quantities adjusting more fully only as the supply side catches up. The resulting short-run gap between demand and supply comprises the "race between education and technology" (Goldin and Katz, 2007), creating cohort-specific wage differentials reflecting imperfect substitution of skills across workers (Adão et al., 2024). This leaves an open question for future research.

5. Results

5.1. Validation: Disclosure and Job Post Mentions of Generative AI

Table 2 validates that prior investment in data integration predicts GAI adoption using a GAI-specific keyword dictionary designed to minimize false positives. All models use cross-sectional regressions relating 2022 data integration to cumulative post-ChatGPT GAI mentions, with industry fixed effects and 2022 controls. Panel A examines GAI mentions across four data sources spanning both investor-facing disclosures and labor market outcomes. A one-standard-deviation increase in data integration is associated with a 4% increase in 10-K GAI mentions, a 22% increase in earnings call mentions, a 16% increase in job posting mentions across all occupations, and an 11% increase in mentions in the résumés of employees. The consistency across sources addresses concerns that firms might over-represent AI capabilities in investor-facing disclosures (Barrios et al., 2025), supporting the interpretation that data integration captures genuine differences in GAI exposure rather than disclosure behavior.

However, Panel A by itself does not necessarily suggest that firms are adopting GAI in their accounting function. To examine this Panel B disaggregates the job posting and résumé results by occupation to assess whether data-integrated firms see greater GAI mentions within the accounting function specifically, rather than just at the firm level. For job postings, a

one-standard-deviation increase in prior investment in data integration corresponds to an 11% increase in core accounting GAI mentions and a 16% increase for the broader accounting and finance function. For employee résumés, the corresponding increases are 10% and 18%. This occupation-specific evidence reinforces that data integration drives GAI adoption in functions like accounting where integration enables the technology.

5.2. Employment and Labor Demand

Table 3 examines the effect of data integration on accounting employment and labor demand. Panel A estimates difference-in-differences regressions of staff-level employee counts on data integration with firm and year fixed effects and controls. Columns (1) and (3) report pooled post-treatment effects: a one-standard-deviation increase in data integration is associated with a 2.7% decrease in core accounting employees and a 2.3% decrease in the broader accounting and finance function. To put this in context, the mean firm in the sample employs approximately 48 staff-level accountants (Table 1: 12 staff-level accountants for the median), so a 2.7% reduction corresponds to about 1.3 accountants. Columns (2) and (4) decompose by year: effects are statistically insignificant in 2023, strengthen to 2.2–2.6% in 2024, and reach 4.3–4.7% by 2025—implying a reduction of approximately 2 staff accountants at the mean firm by the end of the sample period. Figure 2 plots the corresponding event study coefficients, confirming flat pre-trends and a sharp post-2022 divergence. This gradual pattern is consistent with firms progressively adopting and integrating GAI tools rather than implementing immediate workforce reductions.

Panel B examines job postings, which provide a forward-looking measure of labor demand; the firm-year posting-count panel begins in 2016, when Revelio's job posting coverage becomes substantial. Columns (1) and (3) estimate difference-in-differences with firm and year fixed effects: a one-standard-deviation increase in data integration corresponds to a 13% reduction in core accounting postings and an 8% reduction in broader accounting postings. These magnitudes are larger than the contemporaneous employment effects in Panel A,

consistent with job postings capturing intended future hiring reductions before they fully materialize in headcount (Davis et al., 2013). Columns (2) and (4) apply the within-firm accountant-versus-legal triple-difference to job postings: the relative decline is statistically significant for the broader accounting and finance function but directionally consistent and statistically insignificant for the core occupation, indicating that the posting-based within-firm contrast is noisier than the corresponding employment comparison.

Panel C tests whether the employment reductions are accounting-specific using a within-firm triple-difference that compares accountants to legal employees at the same firm in the same year. Legal employees serve as a natural comparison: they are similarly credentialed professionals (Bloomfield et al., 2017; Cascino et al., 2021), but legal AI tools rely less on data integration, operating primarily on static documents rather than dynamic operational data. The data are stacked to two rows per firm-year (one accounting, one legal) with firm×year, firm×occupation, and year×occupation fixed effects, isolating the differential effect of data integration on accounting versus legal headcount. A one-standard-deviation increase in data integration is associated with a 12–13% relative decline in accounting employment compared to legal. The year-by-year decomposition shows that this differential is present from 2023, when accounting employment falls roughly 9% relative to legal, and strengthens to about 14% by 2025. The analogous IT-intensity placebo interaction enters positively rather than negatively, reinforcing that the effect operates through data integration specifically rather than general technological sophistication. Figure 4 plots the event study, and Table OA.2 confirms statistically insignificant effects for legal employees in a specification parallel to Panel A.

5.3. The Hiring Channel

Having documented a decline in staff-level accounting headcount, I next ask whether it operates through reduced hiring or increased separations. Table 4 decomposes the staff-level employment decline into its hiring and attrition components using firm-level employee flow data constructed from Revelio. An entrant is defined as an employee whose first observed

start date at a firm falls within a given year; a separation is defined as an employee whose final observed end date at the firm falls within the year. Internal role changes within the same firm are not counted, ensuring that these flows capture true firm entry and exit rather than internal transfers. Because constructing year-end separations requires a full subsequent year of data, this analysis runs through 2024.

Panel A estimates changes in the number of staff-level new hires entering the firm. A one-standard-deviation increase in data integration is associated with an 8% reduction in core accounting new hires and an 8.5% reduction in the broader function, both significant at the 1% level. The year-by-year decomposition shows the hiring decline present and significant in both 2023 and 2024. Panel B examines separations. In contrast to the sharp hiring decline, I find no significant change in the rate of staff-level separations at data-integrated firms across either occupation definition. This asymmetry between hiring and attrition provides direct evidence on the mechanism: the employment reductions documented in Panel A of Table 3 are driven by reduced hiring rather than increased turnover or layoffs. Data-integrated firms appear to be allowing natural attrition to reduce headcount while curbing replacement hiring. Figure 5 overlays the two event studies, confirming a post-2022 divergence in hiring with no comparable change in separations.

5.4. Composition and Wages

I then investigate compositional shifts within the accounting function of data-integrated firms in the post-GAI period. GAI reallocates the assignment of tasks between labor and capital, and the reallocation is likely strongest in roles that are more routine in nature—typically staff-level junior positions—while manager-level employees may experience greater augmentation as their roles transition to a more strategic capacity.

Table 5 tests this using a within-firm triple-difference design that compares manager- and staff-level accountants in the same firm-year. Panel A stacks the data to two rows per firm-year (one manager, one staff) with firm×year, firm×role, and year×role fixed effects. Columns (1)

and (3) show that a one-standard-deviation increase in prior investment in data integration is associated with a 3.6% relative increase in manager employment compared to staff for core accounting and a 2.7% increase for the broader function. Columns (2) and (4) decompose by year: the differential is present from 2023 and strengthens over time, reaching 5.7% (3.9%) by 2025. Figure 3 plots the event study for manager-level employment in the baseline DiD specification (parallel to Panel A of Table 3), confirming flat and statistically insignificant effects across all post-treatment years; the full-window event study is presented in Figure OA.1 of the online appendix. Combined with the staff-level declines documented in Table 3, these results indicate that the employment reductions at data-integrated firms concentrate in junior staff positions while managerial headcount is stable.

Panel B turns to posted wages, restricting to postings with disclosed (non-estimated) salaries. Staff-level wage effects are economically small and statistically insignificant. In contrast, manager-level postings at data-integrated firms show significant wage increases: a one-standard-deviation increase in data integration corresponds to a 4–6% increase in posted salary, robust to the inclusion of controls for the task content of the job. This pattern—higher wages for managers but not staff—is consistent with Autor and Thompson (2025): as GAI automates routine tasks, the remaining senior work becomes more valuable because it requires greater expertise.

5.5. Robustness

To address concerns that my results are sensitive to specific research design choices, I implement several robustness tests in Table 6, with inferences unchanged across specifications. A key concern with the baseline results is that investment in data integration may always reduce accounting headcount—that is, the employment effects reflect a consequence of integration investment rather than a GAI-specific channel. I address this and other identification concerns with several sets of alternative specifications and measurement choices.

Panel A presents robustness tests for the staff-level employment DiD, each shown in

pooled and event-time form. Columns (1)–(2) re-weight the sample using entropy balancing on observable covariates, reducing concerns that selection into data integration drives my findings. Columns (3)–(4) replace the 2022-vintage treatment with a 2019 base year, measuring data integration exposure prior to the pandemic. If the results reflected a permanent main effect of integration investment, the 2019 measure should predict similar employment declines in the pre-ChatGPT period—but event-time estimates show no effect before 2023, with the divergence emerging only after ChatGPT's release. Columns (5)–(6) replace the broad data integration measure with a narrower measure capturing the share of employee résumés mentioning specific data integration software; this proxy yields a directionally consistent but weaker effect that reaches significance only by 2025. Figure OA.2 plots the corresponding event study coefficients, confirming flat pre-trends and post-2022 divergence under the principal alternative specifications.

Panel B replicates the manager-versus-staff triple-difference under four alternative treatments: the baseline 2022 measure, the 2019 base year, DI Software Intensity, and a three-year rolling average of data integration investment (2020–2022). The positive and significant differential favoring managers is robust across all specifications, reinforcing that GAI-driven automation disproportionately affects staff-level positions.

Table OA.4 in the online appendix presents additional robustness checks for the baseline staff employment DiD: absorbing ERP vendor $\times$ year fixed effects to control for differential trends within enterprise software ecosystems, replacing year fixed effects with industry $\times$ year fixed effects to absorb industry-specific shocks, estimating via OLS on the inverse hyperbolic sine transformation, and an IT-only placebo that omits data integration entirely. All yield consistent inferences. Finally, Table OA.1 reports placebo tests that re-estimate the baseline DiD at two earlier treatment windows (2019 and 2016) where no GAI-specific effect should be present. Neither placebo produces significant effects, providing direct evidence against the concern that data integration has a permanent main effect on accounting employment independent of GAI.

5.6. The Task Content of Accounting Jobs

The results so far validate the use of data integration as an instrument and then use that instrument to document a reduction in labor demand for staff-level accountants, driven by reduced hiring, and a compositional shift in demand toward manager-level roles. However, these results do not speak to how the nature of remaining accounting work is likely to change as a result of GAI. It is possible that fewer staff accountants are needed but that these remaining staff accountants are expected to perform the same work as before and are simply more productive at performing this work. On the other hand, it is possible that these accountants will be reallocated to higher-value work and expected to perform different tasks.

The remainder of my analyses use the text of job postings and résumé data first to examine changes in the task content of accounting jobs, holding seniority constant, and then to investigate the consequences of this changing task content. To focus on a consistent mechanism, these analyses focus specifically on the job posts and résumés for employees in the core accounting function. First, I examine changes in the task content of accounting jobs. To do so, I use the dictionary of 29 tasks listed for accountants and auditors from the O*NET taxonomy provided by the US Department of Labor. I feed the text of each job posting through a large language model and generate 29 separate binaries representing whether each of the 29 accounting job tasks is mentioned in a given posting. I then group and sum these binaries into three distinct categories: routine/compliance, analytic, and strategic tasks. All tasks that do not clearly fit into one of these three categories comprise the residual and are only used as a control in the remaining analysis. The specific task groupings are included in the online appendix.

Table 7 uses this classification to examine changes in task content at the firm-year level by aggregating job postings to firm-year means. Columns (1)–(3) examine the share of routine, analytic, and strategic tasks relative to total tasks. Accounting job postings at firms with greater prior investment in data integration are associated with a statistically

significant reduction in routine task intensity. Analytic and strategic task shares show positive but individually insignificant point estimates. Columns (4) and (5) examine relative ratios—analytic-to-routine and strategic-to-routine—and find statistically significant increases in both. Together, these results indicate that data-integrated firms are recomposing accounting job requirements away from routine work and toward analytic and strategic work, consistent with GAI substituting for routine tasks while complementing higher-order judgment.

5.7. Skill Requirements: CPA Demand and New Hires

As firms shift their accountants toward more analytical and strategic work, different human capital is likely to be complementary to the nature of remaining accounting work. Table 8 examines how data integration affects the demand for CPA certification in job postings and the characteristics of the accountants that data-integrated firms actually hire.

Panel A studies the likelihood that a core accounting job posting requires or prefers a CPA. CPA certification provides a standardized benchmark of accounting expertise, making it a natural outcome to track how technological change alters the skill level expected from accountants. Columns (1) and (2) focus on staff-level postings: a one-standard-deviation increase in data integration is associated with a 0.9–1.0 percentage point increase in the likelihood of requiring a CPA. Given a sample mean CPA requirement rate of approximately 44% (Table 1), this corresponds to a relative increase of about 2%. Columns (3) and (4) show that the effect is concentrated entirely in staff-level roles: manager-level postings exhibit no significant change in CPA requirements, consistent with the interpretation that data-integrated firms are raising the credential bar specifically for the junior positions that remain after automation displaces routine work.

Panel B uses résumé text data to examine the characteristics of the accountants who *are* hired at data-integrated firms in the post-GAI period. New hires are significantly more likely to hold a CPA, an accounting degree, a technical degree, a master's degree, and a Master of Accounting. These results indicate that while data-integrated firms hire fewer accountants

overall—as documented in Table 4—the accountants they do hire are more credentialed, consistent with a rising skill threshold for entry into the profession. The increased prevalence of technical degrees alongside traditional accounting credentials suggests demand for hybrid skill sets that combine domain expertise with analytical or computational capabilities.

6. Conclusion

This paper provides early large-sample evidence on how generative AI is reshaping the market for corporate accountants. Because the returns to a general-purpose technology depend on complementary investments, I focus on data integration—the infrastructure that makes organizational knowledge accessible to GAI within accounting workflows—and exploit within-industry variation in firms' prior investment. Firms with greater pre-existing data integration adopt GAI more intensively across disclosures, job postings, and employee résumés, and subsequently reorganize their accounting workforce.

The reorganization is sharp but selective. Data-integrated firms reduce staff-level accounting employment by 2.7% per standard deviation of integration, with effects growing through 2025. The adjustment operates almost entirely through a reduction in hiring rather than changes in firing: new entrants fall by 8% while separations are unchanged, indicating that firms shrink the junior ranks through attrition rather than layoffs. Managerial headcount is stable and rises relative to staff, and new manager-level roles command wage premia of 4–6%. GAI, for now, is not eliminating the accounting function but rather hollowing out its entry point. The work that remains is changing in tandem. Job postings at data-integrated firms shift away from routine tasks and toward analytic and strategic work, staff-level postings increasingly require CPA certification, and the accountants these firms *do* hire are more likely to hold accounting degrees, advanced degrees, and technical training. Fewer junior accountants are being hired, and those who are face a higher bar: more credentials, more complex work, and a steeper path to the expertise the remaining tasks demand.

Together, these results reconcile why GAI can demonstrably raise productivity in accounting tasks, yet aggregate studies find little change in employment (Choi and Xie, 2026; Humlum and Vestergaard, 2025a). The early effects are concentrated among firms whose data infrastructure permits enterprise-level adoption, and within those firms, among the most junior and routine-intensive roles. Technological shocks are mediated by organizational infrastructure, and the internal information environment determines not just whether firms adopt GAI but who bears its labor market consequences. These findings carry implications moving forward: if the profession's traditional apprenticeship model—in which junior staff learn by performing the routine work that GAI now absorbs—erodes faster than training and credentialing adjust, the supply of experienced accountants may tighten precisely as the expertise premium rises. How educators, firms, and the profession respond to this race between technology and human capital formation is an important question for future research.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author(s) used Claude Code, Claude, Codex, and ChatGPT in order to minimize errors in preparing the data for analysis and to improve the readability and language of the manuscript. The author(s) reviewed and edited the output as needed and take full responsibility for the content of the published article.

References

Abis, S., Veldkamp, L., 2024. The changing economics of knowledge production. *The Review of Financial Studies* 37, 89–118.

Acemoglu, D., Autor, D., 2011. Skills, tasks and technologies: Implications for employment and earnings. In: *Handbook of labor economics*, Elsevier, vol. 4, pp. 1043–1171.

Acemoglu, D., Restrepo, P., 2018. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review* 108, 1488–1542.

Adão, R., Beraja, M., Pandalai-Nayar, N., 2024. Fast and slow technological transitions. *Journal of Political Economy Macroeconomics* 2, 183–227.

Andrusko, M., Amble, S., 2024. Death, taxes, and ai: How generative ai will change accounting. Andreessen Horowitz (a16z) Blog, posted June 24, 2024.

Ashraf, M., 2025. Does automation improve financial reporting? evidence from internal controls. *Review of Accounting Studies* 30, 436–479.

Autor, D., Thompson, N., 2025. Expertise. *Journal of the European Economic Association* 23, 1203–1271.

Autor, D. H., Katz, L. F., Krueger, A. B., 1998. Computing inequality: have computers changed the labor market? *The Quarterly Journal of Economics* 113, 1169–1213.

Autor, D. H., Levy, F., Murnane, R. J., 2003. The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics* 118, 1279–1333.

Barrios, J. M., Campbell, J. L., Johnson, R. G., Liu, Y. C., 2025. Artificially intelligent or artificially inflated? determinants and informativeness of corporate ai disclosures. Available at SSRN 5133107 .

Berger, P. G., Cai, W., Qiu, L., Shen, C. X., 2024. Employer and employee responses to generative ai: Early evidence. Available at SSRN 4874061 .

Bessen, J., 2019. Automation and jobs: When technology boosts employment. Economic

Policy 34, 589–626.

Bick, A., Blandin, A., Deming, D. J., 2026. The rapid adoption of generative ai. Management Science .

Blankespoor, E., Dehaan, E., Li, Q., 2026. Generative ai in financial reporting. Journal of Accounting Research .

Bloom, N., Garicano, L., Sadun, R., Van Reenen, J., 2014. The distinct effects of information technology and communication technology on firm organization. *Management Science* 60, 2859–2885.

Bloomfield, M. J., Brüggemann, U., Christensen, H. B., Leuz, C., 2017. The effect of regulatory harmonization on cross-border labor migration: Evidence from the accounting profession. *Journal of Accounting Research* 55, 35–78.

Böke, J., De la Parra, D., Gallemore, J., Glaeser, S., 2024. Artificial intelligence and white-collar work: Evidence from the accounting profession. Available at SSRN 5039640 .

Bresnahan, T. F., Brynjolfsson, E., Hitt, L. M., 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics* 117, 339–376.

Bresnahan, T. F., Trajtenberg, M., 1995. General purpose technologies 'engines of growth'? *Journal of Econometrics* 65, 83–108.

Brynjolfsson, E., Chandar, B., Chen, R., 2025a. Canaries in the coal mine?: Six facts about the recent employment effects of artificial intelligence .

Brynjolfsson, E., Li, D., Raymond, L., 2025b. Generative ai at work. *The Quarterly Journal of Economics* 140, 889–942.

Bughin, J., Lacroix, M., Luna, P., Nonyane, M., Purcell, E., Restrepo, M., van Kuiken, S., 2023. The data dividend: Fueling generative ai. Accessed: 2025-08-01.

Callaway, B., Goodman-Bacon, A., Sant'Anna, P. H., 2024. Difference-in-differences with a continuous treatment. Tech. rep., National Bureau of Economic Research.

Card, D., Lemieux, T., 2001. Can falling supply explain the rising return to college for younger

men? a cohort-based analysis. *The Quarterly Journal of Economics* 116, 705–746.

Cascino, S., Tamayo, A., Vetter, F., 2021. Labor market effects of spatial licensing requirements: Evidence from cpa mobility. *Journal of Accounting Research* 59, 111–161.

Challapally, A., Pease, C., Raskar, R., Chari, P., 2025. The genai divide: State of ai in business 2025. Tech. rep., MIT NANDA, preliminary findings from AI implementation research.

Charoenwong, B., Kowaleski, Z. T., Kwan, A., Sutherland, A. G., 2024. Regtech: Technology-driven compliance and its effects on profitability, operations, and market structure. *Journal of Financial Economics* 154, 103792.

Choi, B.-G., Choi, J. H., McNichols, M. F., Zhou, F., 2024. The market for accountants. Available at SSRN 5101847 .

Choi, J. H., Xie, C. L., 2026. Human+ ai in accounting: Early evidence from the field. Journal of Accounting Research 64, 1333–1373.

Cohn, J. B., Liu, Z., Wardlaw, M. I., 2022. Count (and count-like) data in finance. *Journal of Financial Economics* 146, 529–551.

Davenport, T., Tiwari, P., 2024. Is your company's data ready for generative ai. *Harvard Business Review* 5.

Davis, S. J., Faberman, R. J., Haltiwanger, J. C., 2013. The establishment-level behavior of vacancies and hiring. *The Quarterly Journal of Economics* 128, 581–622.

Deming, D., Silliman, M., 2025. Skills and human capital in the labor market. In: *Handbook of Labor Economics*, Elsevier, vol. 6, pp. 115–157.

Demirer, M., Horton, J. J., Immorlica, N., Lucier, B., Shahidi, P., 2026. Chaining tasks, redefining work: A theory of ai automation. Tech. rep., National Bureau of Economic Research.

Dorantes, C.-A., Li, C., Peters, G. F., Richardson, V. J., 2013. The effect of enterprise systems implementation on the firm information environment. *Contemporary Accounting Research* 30, 1427–1461.

Eisfeldt, A. L., Schubert, G., Zhang, M. B., 2023. Generative ai and firm values. Tech. rep.,

National Bureau of Economic Research.

Eloundou, T., Manning, S., Mishkin, P., Rock, D., 2024. Gpts are gpts: Labor market impact potential of llms. *Science* 384, 1306–1308.

Fedyk, A., Hodson, J., Khimich, N., Fedyk, T., 2022. Is artificial intelligence improving the audit process? *Review of Accounting Studies* 27, 938–985.

Frank, M. R., Sabet, A. J., Simon, L., Bana, S. H., Yu, R., 2026. Ai-exposed jobs deteriorated before chatgpt. arXiv preprint arXiv:2601.02554 .

Freund, L., Mann, L., 2026. Job transformation, specialization, and the labor market effects of ai. Tech. rep., CESifo Working Paper.

Friedman, H. L., Sutherland, A., Vetter, F., 2024. Technological investment and accounting: A demand-side perspective on accounting enrollment declines. Available at SSRN 4707807 .

Gallemore, J., Labro, E., 2015. The importance of the internal information environment for tax avoidance. *Journal of Accounting and Economics* 60, 149–167.

Gans, J. S., Goldfarb, A., 2026. O-ring automation. Tech. rep., National Bureau of Economic Research.

Goldin, C., Katz, L. F., 2007. The race between education and technology: the evolution of us educational wage differentials, 1890 to 2005.

Greenwood, J., Hercowitz, Z., Krusell, P., 1997. Long-run implications of investment-specific technological change. *American Economic Review* pp. 342–362.

Ham, C., Hann, R. N., Rabier, M., Wang, W., 2025. Auditor skill demands and audit quality: Evidence from job postings. *Management Science* 71, 5805–5829.

Hampole, M., Papanikolaou, D., Schmidt, L. D., Seegmiller, B., 2025. Artificial intelligence and the labor market. Tech. rep., National Bureau of Economic Research.

Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., Mueller, J., Hong, J., Ritchie, S., Belonax, T., et al., 2025. Which economic tasks are performed with ai? evidence from millions of claude conversations. arXiv preprint arXiv:2503.04761 .

Heese, J., Pacelli, J., 2025. Enterprise resource planning (erp) system implementations and

corporate misconduct. *The Accounting Review* 100, 291–315.

Helpman, E., Trajtenberg, M., 1994. A time to sow and a time to reap: Growth based on general purpose technologies.

Humlum, A., Vestergaard, E., 2025a. Large language models, small labor market effects. Tech. rep., National Bureau of Economic Research.

Humlum, A., Vestergaard, E., 2025b. The unequal adoption of chatgpt exacerbates existing inequalities among workers. *Proceedings of the National Academy of Sciences* 122, e2414972121.

Katz, L. F., Murphy, K. M., 1992. Changes in relative wages, 1963–1987: supply and demand factors. *The Quarterly Journal of Economics* 107, 35–78.

Kim, D. S., Sutherland, A., Vetter, F., 2025. Software adoption through the lens of directed technological change. Available at SSRN 5369069 .

Kogan, L., Papanikolaou, D., Schmidt, L. D., Seegmiller, B., 2023. Technology and labor displacement: Evidence from linking patents with worker-level data. Tech. rep., National Bureau of Economic Research.

KPMG, 2024. Beyond the hype: Genai is delivering, new KPMG survey shows Accessed on May 7, 2025.

Law, K. K., Shen, M., 2024. How does artificial intelligence shape audit firms? *Management Science* .

Lichtinger, G., Hosseini Maasoum, S. M., 2025. Generative ai as seniority-biased technological change: Evidence from us resume and job posting data. Available at SSRN 5425555 .

Lodefalk, M., Löthman, L., Koch, M., Engberg, E., 2026. Same storm, different boats: Generative ai and the age gradient in hiring. Tech. rep., Working Paper.

McAfee, A., Rock, D., Brynjolfsson, E., 2023. How to capitalize on generative ai. *Harvard Business Review* 101, 42–48.

Morris, J. J., 2011. The impact of enterprise resource planning (erp) systems on the effectiveness of internal controls over financial reporting. *Journal of Information Systems* 25,

129–157.

Noy, S., Zhang, W., 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381, 187–192.

Salesforce, 2025. 2025 connectivity benchmark report. Tech. rep., MuleSoft, a Salesforce company, accessed: 2025-08-20.

Schubert, G., 2025. Organizational technology ladders: Remote work and generative ai adoption. Available at SSRN 5094265 .

Solow, R., 1987. We'd better watch out. *New York Times Book Review* 36.

Valentine, M., Politzer, D. J., Davenport, T. H., 2025. How to make enterprise gen AI work. *Harvard Business Review* .

7. Appendix

Appendix A: Variable Definitions

Note: All continuous variables winsorized at 1% on each tail unless otherwise noted.

Variable	Definition	Source
Data Integration$_{i,2022}$	Sum of 2022 Revelio-predicted salaries (using Revelio employment weights) paid to data management workers (O*NET 15-1242 and 15-1243) divided by sum of 2022 Revelio-predicted salaries paid to all white-collar workers (O*NET 11-29, 41 and 43) at firm i in year 2022. De-meaned by 3-digit NAICS industry average.	Revelio
Unadjusted DataIntegration$_{i,2022}$	Sum of 2022 Revelio-predicted salaries (using Revelio employment weights) paid to data management workers divided by sum of 2022 Revelio-predicted salaries paid to all white-collar workers (O*NET 11-29, 41 and 43) at firm i in year 2022.	Revelio
IT Intensity$_{i,2022}$	Sum of 2022 Revelio-predicted salaries (using Revelio employment weights) paid to non-data IT workers (O*NET 15-1211, 15-1212, 15-1231, 15-1232, 15-1241, 15-1244) divided by sum of 2022 Revelio-predicted salaries paid to all white-collar workers (O*NET 11-29, 41 and 43) at firm i in year 2022. De-meaned by 3-digit NAICS industry average.	Revelio
#AcctEmp (Staff)$_{it}$	Sum of employees classified as non-managers (Revelio seniority code ≤ 3) under O*NET code 13-2011 (Accountants and Auditors) for firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio
#AcctEmp (Manager)$_{it}$	Sum of employees classified as managers (Revelio seniority code > 3) under O*NET code 13-2011 (Accountants and Auditors) for firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio

Appendix A: Variable Definitions (Continued)

Variable	Definition	Source
#AcctPosts (Staff)$_{it}$	Count of unique job posts for O*NET code 13-2011 for firm i in year t with seniority level of 3 or below (non-manager).	Revelio
#AcctPosts (Manager)$_{it}$	Count of unique job posts for O*NET code 13-2011 for firm i in year t with seniority level of 4 or above (manager roles).	Revelio
Firm Age$_{it}$	Number of years between firm's first appearance on CRSP and current year for firm i in year t	CRSP
LnAssets$_{it}$	Natural log of total assets (at) in millions for firm i in year t	Compustat
Capex Scaled$_{it}$	Capital expenditures (capx) scaled by lagged total assets for firm i in year t	Compustat
Leverage$_{it}$	Total debt scaled by total shareholder's equity for firm i in year t	Compustat
ROA$_{it}$	Return on assets. Net income scaled by lagged total assets for firm i in year t.	Compustat
Book-to-Market$_{it}$	Book value of equity at fiscal year end divided by market capitalization at fiscal year end.	Compustat
AverageReturns$_{it}$	Average of daily returns for firm i in year t after handling missing values.	CRSP
ReportingLag$_{it}$	Quarterly data. Count of number of days between fiscal year end and SEC EDGAR filing date for 10-K or 10-Q for firm i in quarter t.	WRDS SEC Analytics
D(Guidance)$_{it}$	Indicator variable equal to 1 for firm i in quarter t if they have issued any form of guidance, 0 otherwise.	IBES
Forecast Accuracy$_{it}$	Mean absolute percentage error for EPS and earnings forecasts issued by firm i in quarter t. This requires that a forecast is issued and realized to calculate the percentage error.	IBES
D(AnyGAI - 10K)$_{it}$	Indicator variable equal to 1 for firm i in year t if they have mentioned any keywords associated with generative AI (see online appendix) in Item 1A of the 10-K.	SEC EDGAR
NumGAI(10K)$_{it}$	Count of generative AI keywords (see online appendix) in Item 1A of the 10-K.	SEC EDGAR
D(AnyGAI - EC)$_{it}$	Indicator variable equal to 1 for firm i in year t if they have mentioned any keywords associated with generative AI (see online appendix) in the presenter speech section of the earnings call.	Capital IQ
NumGAI(EC)$_{it}$	Count of generative AI keywords (see online appendix) in the presenter speech section of the earnings call.	Capital IQ

Appendix A: Variable Definitions (Continued)

Variable	Definition	Source
D(AnyGAI - Job Posts)$_{it}$	Indicator variable equal to 1 for firm i in year t if they have mentioned any keywords associated with generative AI (see online appendix) in any job post listed that year.	Revelio
PostGPT$_t$	Indicator equal to 1 for observations in 2023, 2024, and 2025, after the release of ChatGPT. Observations for December of 2022 fiscal year end classified as 0.	
Acct MgrToStaff EmpRatio$_{it}$	The ratio of the sum of employees classified as managers (Revelio seniority code ≥ 4) under O*NET code 13-2011 (Accountants and Auditors) for firm i in year t to the sum of employees classified as non-managers (Revelio seniority code ≤ 3) under the same code.	Revelio
SalesGrowth$_{it}$	Year-over-year percentage change in revenue for firm i in year t (compared to year t minus 1).	Compustat
#RoutineTasks$_{it}$	Average number of O*NET tasks for code 13-2011 (accountants and auditors) classified as routine (see online appendix) across all core accounting job postings at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
#AnalyticTasks$_{it}$	Average number of O*NET tasks for code 13-2011 (accountants and auditors) classified as analytic (see online appendix) across all core accounting job postings at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
#StrategicTasks$_{it}$	Average number of O*NET tasks for code 13-2011 (accountants and auditors) classified as strategic (see online appendix) across all core accounting job postings at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
#OtherTasks$_{it}$	Average number of O*NET tasks for code 13-2011 (accountants and auditors) not classified as routine, analytic, or strategic (see online appendix) across all core accounting job postings at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
CPARequired$_{ijt}$	Variable for whether a CPA is required based on a large language model classification for accounting job posting j at firm i in year t. Values can be 0 (no CPA mentioned), 0.5 (CPA listed as preferred), or 1 (CPA listed as required).	Revelio, GPT-4.1-Mini
LnSalary$_{ijt}$	Natural log of posted midpoint of salary for any accounting job posting j in year t at firm i with posted wages. This requires disclosed salary information, not estimated data from Revelio.	Revelio

Appendix A: Variable Definitions (Continued)

Variable	Definition	Source
DISoftware Intensity$_{i,2022}$	Natural log of 1 plus the number of resume keywords from all employees at firm i in year 2022 that mention ETL/Integration tools (Informatica, MuleSoft, Airflow, etc.), divided by log of assets in 2022.	Revelio, Compustat
RollingAverage DataIntegration$_{i,2022}$	Average of the sum of Revelio-predicted salaries (using Revelio employment weights) paid to data management workers (O*NET 15-1242 and 15-1243) divided by sum of Revelio-predicted salaries paid to all white-collar workers (O*NET 11-29, 41 and 43) at firm i in years 2018-2022, de-meaned by 3-digit NAICS industry.	Revelio, Compustat
Data Integration$_{i,2016}$	Sum of 2016 Revelio-predicted salaries (using Revelio employment weights) paid to data management workers (O*NET 15-1242 and 15-1243) divided by sum of 2016 Revelio-predicted salaries paid to all white-collar workers (O*NET 11-29, 41 and 43) at firm i in year 2016. De-meaned by 3-digit NAICS industry average.	Revelio
Data Integration$_{i,2019}$	Sum of 2019 Revelio-predicted salaries (using Revelio employment weights) paid to data management workers (O*NET 15-1242 and 15-1243) divided by sum of 2019 Revelio-predicted salaries paid to all white-collar workers (O*NET 11-29, 41 and 43) at firm i in year 2019. De-meaned by 3-digit NAICS industry average.	Revelio

Appendix A: Variable Definitions (Continued)

Variable	Definition	Source
Hiring and Attrition Flows (Table 4)		
#NewHires (Staff)$_{it}$	Count of staff-level (Revelio seniority code ≤ 3) employees under O*NET code 13-2011 (or the broader accounting and finance function) whose first observed start date at firm i falls within year t. Internal role changes within the same firm are not counted.	Revelio
#Separations (Staff)$_{it}$	Count of staff-level (Revelio seniority code ≤ 3) employees under O*NET code 13-2011 (or the broader accounting and finance function) whose final observed end date at firm i falls within year t. Employees who remain active are not counted. Construction requires a full subsequent year of data, so the series runs through 2024.	Revelio
Task Content Shares and Ratios (Table 7)		
RoutineShare$_{it}$	#RoutineTasks$_{it}$ divided by the total number of O*NET tasks (routine + analytic + strategic + other) identified in accounting job posting j at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
AnalyticShare$_{it}$	#AnalyticTasks$_{it}$ divided by the total number of O*NET tasks identified in accounting job posting j at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
StrategicShare$_{it}$	#StrategicTasks$_{it}$ divided by the total number of O*NET tasks identified in accounting job posting j at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
AnalyticToRoutine$_{it}$	#AnalyticTasks$_{it}$ divided by the sum of #AnalyticTasks$_{it}$ and #RoutineTasks$_{it}$ for accounting job posting j at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
StrategicToRoutine$_{it}$	#StrategicTasks$_{it}$ divided by the sum of #StrategicTasks$_{it}$ and #RoutineTasks$_{it}$ for accounting job posting j at firm i in year t.	Revelio, O*NET, GPT-4.1-Mini
CPA Requirement (Table 8, Panel A)		
CPARequired$_{ijt}$	Ordinal measure of CPA demand in accounting job posting j at firm i in year t, based on a large language model classification of the posting text. Equal to 0 (no CPA mentioned), 0.5 (CPA listed as preferred), or 1 (CPA listed as required).	Revelio, GPT-4.1-Mini

Appendix A: Variable Definitions (Continued)

Variable	Definition	Source
New-Hire Credentials (Table 8, Panel B)		
D(CPA, New Hire)$_{ijt}$	Indicator equal to 1 if new accounting hire j at firm i in year t holds a CPA, identified from the employee's profile name and title; 0 otherwise.	Revelio
D(Acct Degree)$_{ijt}$	Indicator equal to 1 if new accounting hire j at firm i in year t holds an accounting degree; 0 otherwise.	Revelio
D(Tech Degree)$_{ijt}$	Indicator equal to 1 if new accounting hire j at firm i in year t holds a technical degree, defined as a degree in computer science, information technology, mathematics, engineering, or statistics; 0 otherwise.	Revelio
D(Masters)$_{ijt}$	Indicator equal to 1 if new accounting hire j at firm i in year t holds any master's degree; 0 otherwise.	Revelio
D(MAcc)$_{ijt}$	Indicator equal to 1 if new accounting hire j at firm i in year t holds a Master of Accounting degree; 0 otherwise.	Revelio
Placebo Occupation Employment (Tables OA.2 and OA.3)		
#LegalEmp (Staff)$_{it}$	Sum of staff-level (Revelio seniority code ≤ 3) legal employees (lawyers and paralegals) at firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio
#ITEmp (Staff)$_{it}$	Sum of staff-level (Revelio seniority code ≤ 3) non-data IT employees (O*NET 15-1211, 15-1212, 15-1231, 15-1232, 15-1241, 15-1244) at firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio
#HREmp (Staff)$_{it}$	Sum of staff-level (Revelio seniority code ≤ 3) core human resources employees at firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio
#DataEmp (Staff)$_{it}$	Sum of staff-level (Revelio seniority code ≤ 3) data management employees (O*NET 15-1242 and 15-1243) at firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio
#DataEmp (Manager)$_{it}$	Sum of manager-level (Revelio seniority code > 3) data management employees (O*NET 15-1242 and 15-1243) at firm i in year t. Employee is required to be at firm for ≥ 6 months during year.	Revelio
Additional Robustness (Table OA.4)		
D(ERP Adopter)$_{i}$	Indicator equal to 1 if any employee at firm i reports an enterprise resource planning software (e.g., SAP, Oracle, Microsoft Dynamics) on their Revelio profile; 0 otherwise. Used to construct ERP adopter $\times$ year fixed effects.	Revelio

8. Figures

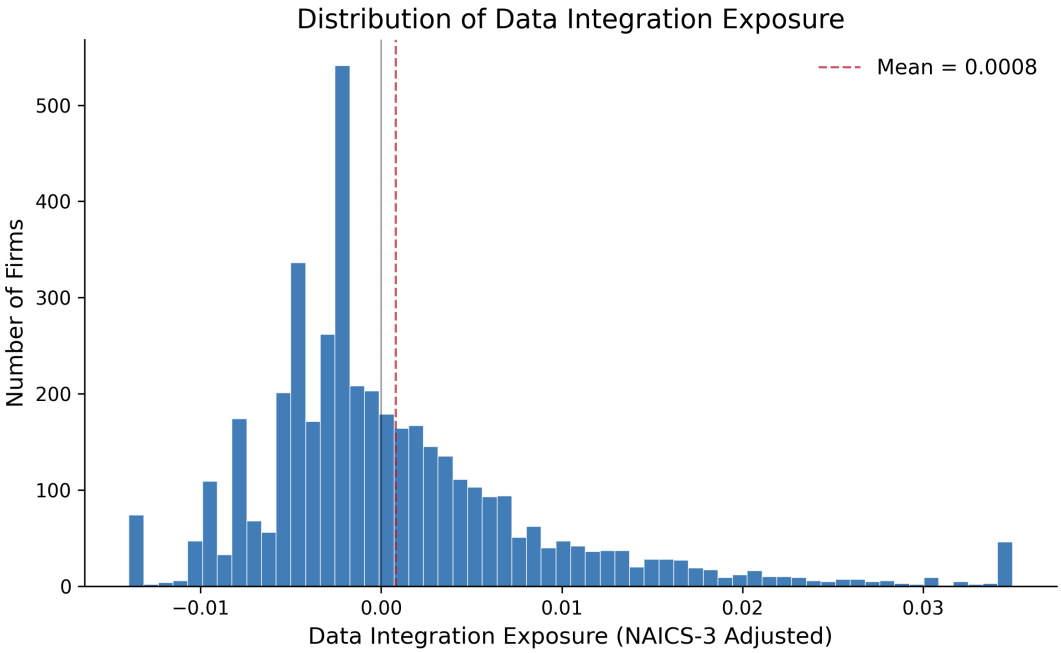


Figure 1. This figure plots the distribution of Data Integration exposure ($DataIntegration_{i,2022}$), using the industry-adjusted (NAICS-3) Data Integration measure for the 2022 cross-section of sample firms, winsorized at the 1st and 99th percentiles. The dashed red line indicates the sample mean.

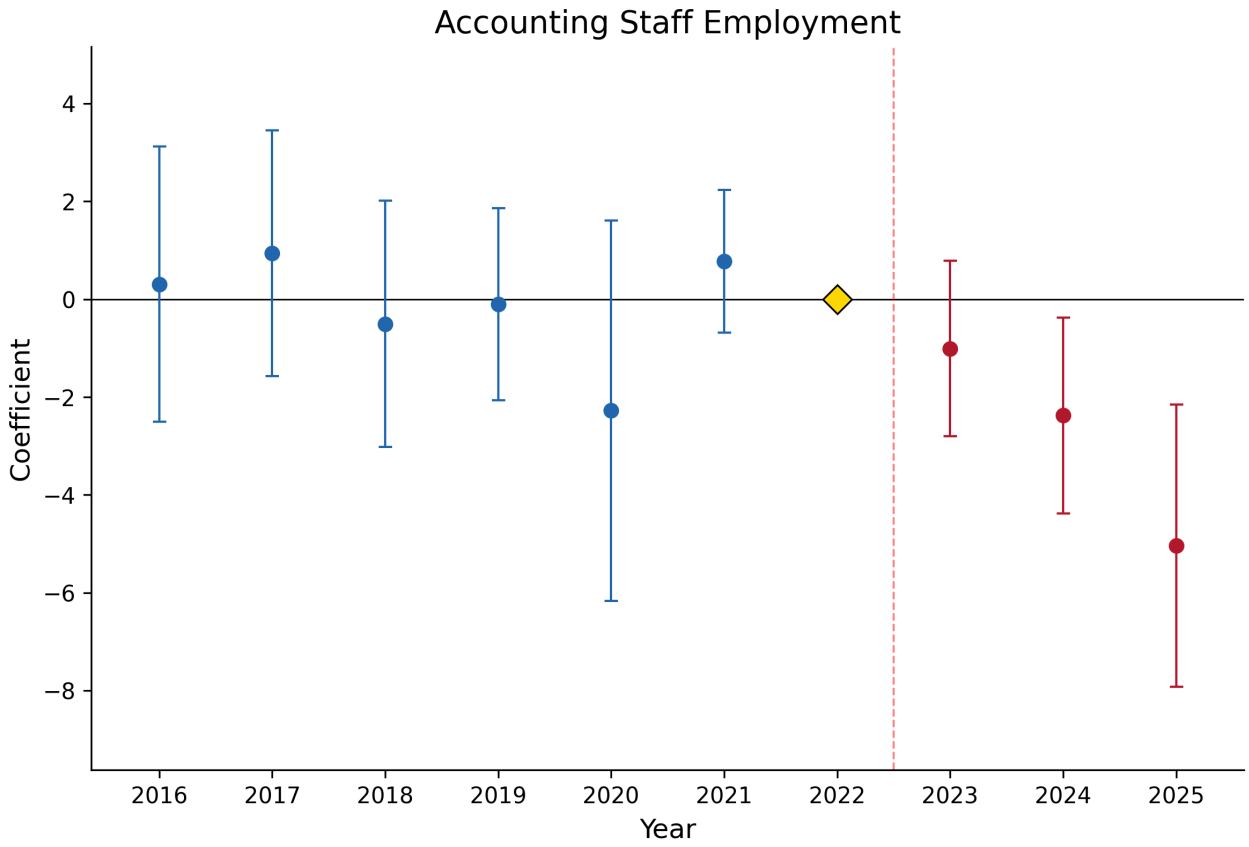


Figure 2. This figure plots coefficient estimates of the effect of data integration on staff-level core accounting employment using an event-time version of Table 3, Panel A, Column 1. The figure plots coefficient estimates and 95% confidence intervals for the interaction between pre-treatment data integration ($DataIntegration_{i,2022}$) and year indicators, with 2022 as the holdout year. The dependent variable is the number of staff-level core accounting employees.

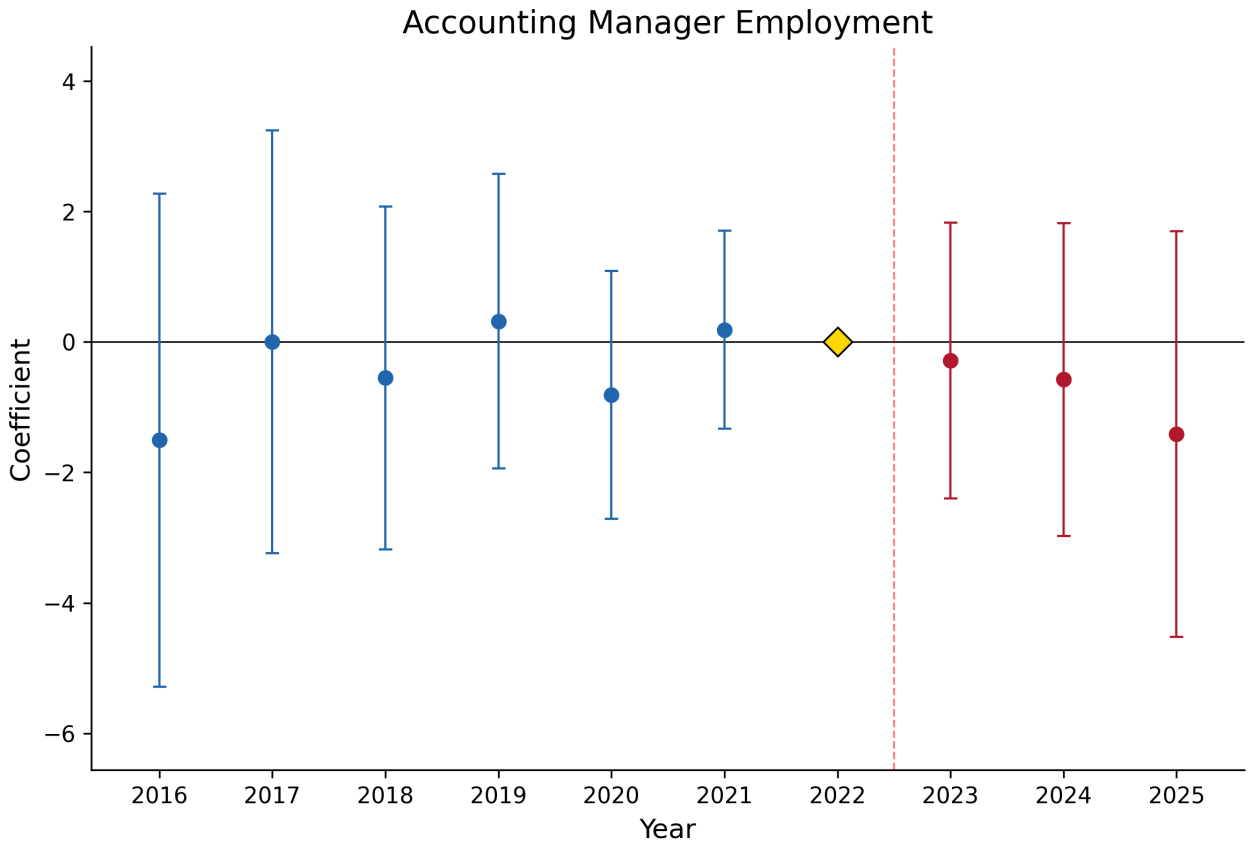


Figure 3. This figure plots coefficient estimates of the effect of data integration on manager-level core accounting employment using an event-time version of Table 3, Panel A, Column 1, but for manager-level employees. Coefficient estimates and 95% confidence intervals for the interaction between pre-treatment data integration ($DataIntegration_{i,2022}$) and year indicators, with 2022 as the holdout year. The dependent variable is the number of manager-level core accounting employees.

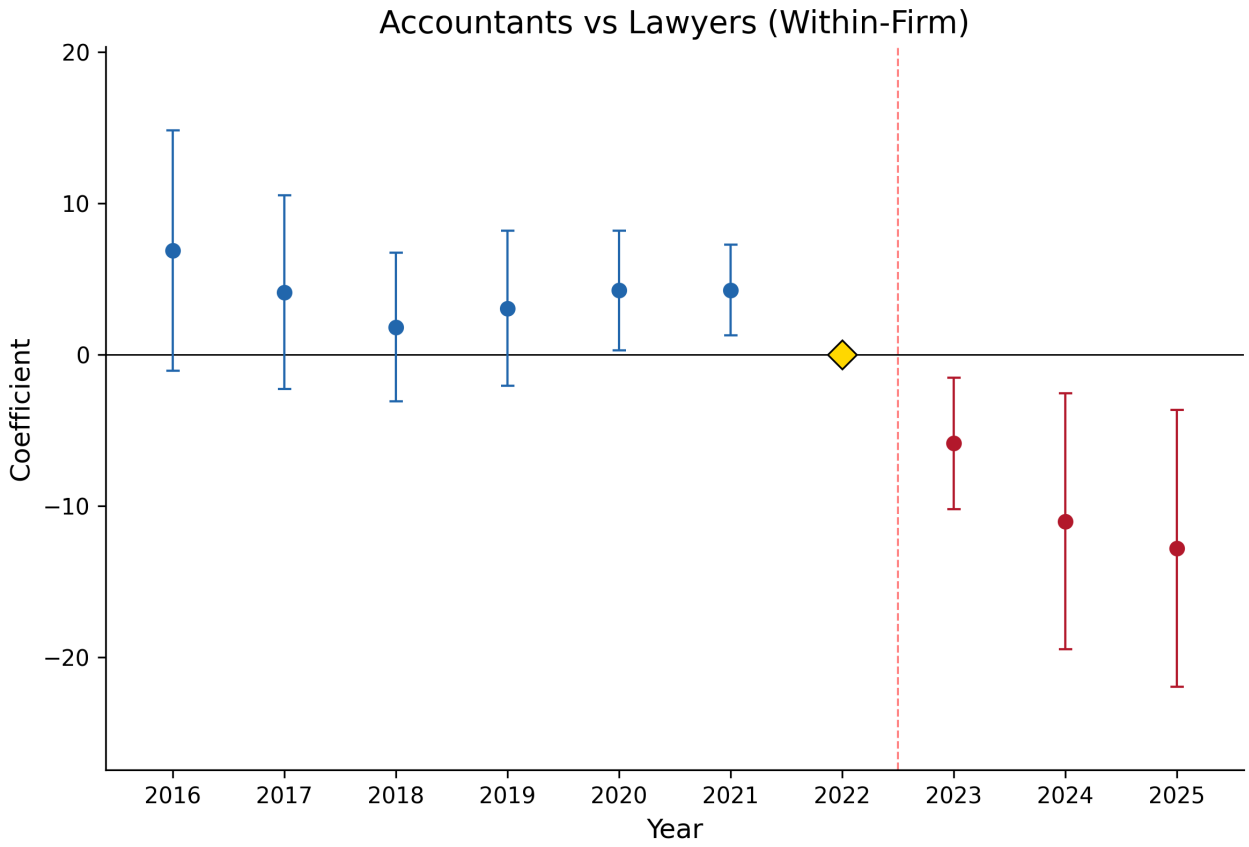


Figure 4. This figure plots coefficient estimates of the within-firm difference-in-differences comparing staff-level core accountants to staff-level lawyers using an event-time version of Table 3, Panel C, Column 1. Coefficient estimates and 95% confidence intervals for the triple interaction between pre-treatment data integration ($DataIntegration_{i,2022}$), an accounting occupation indicator, and year indicators, with 2022 as the holdout year. The dependent variable is the number of staff-level employees in a given cell (core accountants or lawyers).

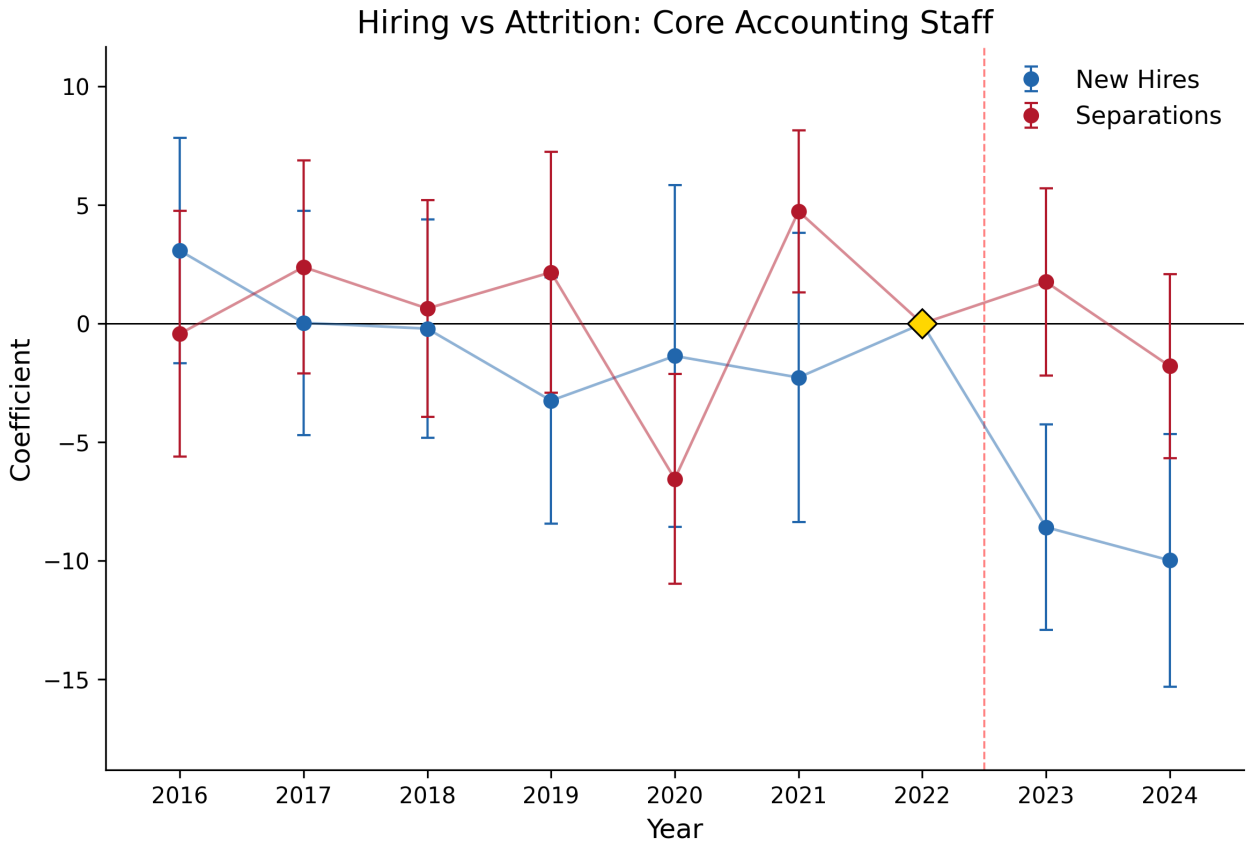


Figure 5. This figure plots coefficient estimates of the effect of data integration on staff-level accounting new hires and separations using an event-time version of Table 4, Panels A and B, Column 1, overlaid. The figure overlays coefficient estimates and 95% confidence intervals from two separate PPML regressions: one for new entrants to the firm (blue) and one for separations (red). Both specifications interact pre-treatment data integration ($DataIntegration_{i,2022}$) with year indicators, with 2022 as the holdout year. The sample spans 2010–2024. The dependent variable is the number of new hires/separations in the core accounting function for a given firm-year.

9. Tables

Table 1. Summary Statistics and Determinants of Data Integration

Panel A: Summary Statistics

	N	Mean	Std. Dev.	P25	Median	P75
Time-Invariant Firm Variables						
Data Integration (NAICS-adj.)$_{i,2022}$	3,620	0.001	0.008	-0.004	-0.001	0.004
Data Integration (unadjusted)$_{i,2022}$	3,620	0.007	0.009	0.000	0.005	0.010
IT Intensity (NAICS-adj.)$_{i,2022}$	3,620	0.001	0.013	-0.004	-0.001	0.004
Time-Varying Firm-Year Variables						
Accounting Employment Share$_{it}$	45,142	0.028	0.037	0.008	0.019	0.033
#Accounting Employees (Staff)$_{it}$	45,142	48.558	101.135	2.157	12.538	44.540
#Accounting Employees (Managers)$_{it}$	45,142	23.720	45.901	1.061	7.190	23.495
Manager-to-Staff Ratio (Acct.)$_{it}$	45,142	0.560	0.575	0.187	0.444	0.750
#Accounting Job Postings (Staff)$_{it}$	45,142	12.555	43.133	0.000	0.000	4.000
Firm Age$_{it}$	45,142	7.272	4.314	3.000	7.000	11.000
Log(Total Assets)$_{it}$	45,142	7.324	2.266	5.825	7.392	8.855
Capital Expenditures / Assets$_{it}$	45,142	0.035	0.049	0.003	0.018	0.046
Return on Assets$_{it}$	45,142	-0.030	0.233	-0.024	0.021	0.069
Sales Growth$_{it}$	45,142	0.106	0.262	-0.031	0.064	0.192
Leverage$_{it}$	45,142	0.262	0.234	0.062	0.218	0.402
Book-to-Market$_{it}$	45,142	0.716	0.704	0.242	0.511	0.915
#Staff-Level Job Postings$_{it}$	12,149	23.4	89.8	2.0	6.0	17.0
#Manager-Level Job Postings$_{it}$	12,149	7.6	37.0	0.0	2.0	5.0
Job Posting Variables (2020–2024)						
Number of Tasks per Posting$_{ijt}$	306,838	8.87	4.44	5.00	8.00	12.00
Routine Task Share$_{ijt}$	306,838	0.443	0.195	0.333	0.417	0.500
Analytic Task Share$_{ijt}$	306,838	0.358	0.167	0.286	0.364	0.438
Strategic Task Share$_{ijt}$	306,838	0.150	0.119	0.000	0.154	0.231
CPA Required or Preferred$_{ijt}$	306,838	0.438	0.496	0.000	0.000	1.000
Staff-Level Posting$_{ijt}$	306,838	0.756	0.430	1.000	1.000	1.000
Manager-Level Posting$_{ijt}$	306,838	0.244	0.430	0.000	0.000	0.000
Ln(Salary), Disclosed Only$_{ijt}$	37,209	11.292	0.495	11.021	11.267	11.549

Panel B: Cross-Sectional Determinants of Data Integration

Dependent Variable:	Data Integration$_{i,2022}$	
	(1)	(2)
Log(Total Assets)	0.001***	0.001***
	(12.548)	(9.673)
Firm Age	-0.000***	-0.000
	(-3.537)	(-0.551)
Capital Expenditures / Assets	-0.005**	-0.005*
	(-2.461)	(-1.779)
Return on Assets	-0.001**	-0.001
	(-2.454)	(-1.582)
Sales Growth	-0.001***	-0.001***
	(-4.325)	(-2.727)
Book-to-Market	-0.001***	-0.000***
	(-4.504)	(-2.605)
Leverage	-0.002***	0.001
	(-2.625)	(1.233)
Number of Business Segments	0.000*	0.000*
	(1.894)	(1.906)
Avg. 10-K Filing Lag (days)	-0.000	-0.000***
	(-1.400)	(-2.869)
Avg. M&A Intensity (aqc/at)	0.015***	-0.003
	(3.219)	(-0.668)
Management Forecast Accuracy	0.099***	0.044
	(2.599)	(1.166)
1(Issues Annual Guidance)	0.002***	0.001
	(3.446)	(1.377)
1(Restatement, fy 2022)	0.001	0.000
	(1.271)	(0.662)
NAICS-3 FEs	No	Yes
Observations	3,620	3,612
R-squared	0.115	0.266

This table reports summary statistics (Panel A) and cross-sectional determinants of firm-level Data Integration (Panel B). Panel A describes the firms, firm-years, and job postings that form the main estimation sample, with statistics computed on the common non-missing subset within each variable group. Job posting statistics are at the individual posting level for core accounting postings (O*NET 13-2011) over 2020–2024. Panel B reports OLS regressions of Data Integration in 2022 on firm characteristics for the 2022 cross-section of firms with complete data on all determinants. Column 1 has no fixed effects; Column 2 includes NAICS-3 industry fixed effects. Robust standard errors; t-statistics in parentheses. Continuous variables are winsorized at the 1st and 99th percentiles; indicators are not winsorized. Detailed variable definitions are provided in the appendix.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 2. Validation: Cross-Sectional GAI Mentions

Panel A: Cross-Sectional Validation (All Occupations)

	10-K	Earnings Calls	Job Postings	Resumes
Dependent Variable:	Number of GAI Mentions			
Data Integration$_{i,2022}$	4.721***	27.662**	20.152***	13.324***
	(3.047)	(2.291)	(3.731)	(3.365)
IT Intensity$_{i,2022}$	3.448***	-0.819	4.735***	4.748***
	(3.589)	(-0.110)	(4.392)	(4.741)
Industry FEs	Yes	Yes	Yes	Yes
Controls (2022)	Yes	Yes	Yes	Yes
Observations	3,620	3,153	3,620	3,620
Pseudo R-squared	0.321	0.328	0.698	0.855

Panel B: Occupation-Specific Validation

	Job Postings		Employee Resumes	
	Core Accounting	Broader Accounting	Core Accounting	Broader Accounting
Dependent Variable:	Number of GAI Mentions			
Data Integration$_{i,2022}$	13.995**	20.137***	12.987*	22.291**
	(2.372)	(2.869)	(1.916)	(2.265)
IT Intensity$_{i,2022}$	4.177***	3.718***	4.904***	2.669
	(3.128)	(3.052)	(4.202)	(1.109)
Industry FEs	Yes	Yes	Yes	Yes
Controls (2022)	Yes	Yes	Yes	Yes
Observations	3,620	3,620	3,620	3,620
Pseudo R-squared	0.438	0.577	0.533	0.715

T-statistics in parentheses. Standard errors are clustered at the firm level. All models use PPML (ppmlhdfe) on cumulative post-ChatGPT (2023–2025) GAI mention counts with 3-digit NAICS industry fixed effects and 2022 controls (firm age, log assets, capital expenditures, ROA, sales growth, leverage, book-to-market, stock returns, employment, and accounting share). Panel A estimates the cross-sectional relationship between industry-adjusted Data Integration and GAI mentions across four data sources. Panel B disaggregates by occupation: "Core Accounting" corresponds to Accountants and Auditors (ONET 13-2011); "Broader Accounting" adds Financial Analysts, Financial Managers, Budget Analysts, and related finance roles. Observation counts vary across columns because ppmlhdfe drops observations with no within-industry variation in the outcome (e.g., industries where no firm has any GAI mentions).

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 3. Effect of Data Integration on Accountant Employment

Panel A: Staff-Level Employees

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of Employees			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	-3.317***		-2.917***	
	(-2.611)		(-2.870)	
$2023_t \times \text{Data Integration}_{i,2022}$		-1.849		-1.572
		(-1.595)		(-1.552)
$2024_t \times \text{Data Integration}_{i,2022}$		-3.224**		-2.781***
		(-2.553)		(-2.639)
$2025_t \times \text{Data Integration}_{i,2022}$		-5.871***		-5.364***
		(-3.527)		(-4.390)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	-1.054	-1.087	-1.493*	-1.536**
	(-1.019)	(-1.049)	(-1.922)	(-1.967)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	40,846	40,846	42,862	42,862
Pseudo R-squared	0.969	0.969	0.985	0.985

Panel B: Job Postings

	Core Accounting		Broader Accounting	
	Staff	Acct vs Legal	Staff	Acct vs Legal
Dependent Variable:	Number of Job Postings			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	-15.842**		-9.885**	
	(-2.218)		(-2.453)	
$\text{PostGPT}_t \times \text{Data Integr.}_{i,2022} \times \text{Acct}_j$		-12.728		-15.553**
		(-1.440)		(-2.233)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	3.624		-0.262	
	(0.915)		(-0.191)	
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022} \times \text{Acct}_j$		-7.417		-4.665
		(-1.112)		(-0.980)
Firm FEs	Yes	–	Yes	–
Year FEs	Yes	–	Yes	–
Firm × Year FEs	–	Yes	–	Yes
Firm × Occ FEs	–	Yes	–	Yes
Year × Occ FEs	–	Yes	–	Yes
Controls	Yes	–	Yes	–
Observations	26,803	39,568	27,649	43,102
Pseudo R-squared	0.965	0.991	0.959	0.993

Panel C: Within-Firm Accountants vs. Legal

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of Employees			
$\text{PostGPT}_t \times \text{Data Integr.}_{i,2022} \times \text{Acct}_j$	-14.804***		-16.798***	
	(-3.148)		(-3.006)	
$2023_t \times \text{Data Integr.}_{i,2022} \times \text{Acct}_j$		-11.037***		-13.280***
		(-3.094)		(-3.042)
$2024_t \times \text{Data Integr.}_{i,2022} \times \text{Acct}_j$		-16.209***		-18.046***
		(-2.935)		(-2.899)
$2025_t \times \text{Data Integr.}_{i,2022} \times \text{Acct}_j$		-18.069***		-19.899***
		(-3.184)		(-2.925)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022} \times \text{Acct}_j$	8.325***	8.234***	8.250***	8.161***
	(2.876)	(2.843)	(2.786)	(2.757)
Firm × Year FEs	Yes	Yes	Yes	Yes
Firm × Occ FEs	Yes	Yes	Yes	Yes
Year × Occ FEs	Yes	Yes	Yes	Yes
Observations	30,950	30,950	31,258	31,258
Pseudo R-squared	0.990	0.990	0.996	0.996

T-statistics in parentheses. Standard errors are clustered at the firm level. Panel A estimates the effect of industry-adjusted (3-digit NAICS) Data Integration on staff-level employee counts using PPML (ppmlhdfe) with firm and year fixed effects and controls. Columns 1 and 3 report pooled post-treatment effects; columns 2 and 4 decompose by year. Panel B examines job posting counts; the job posting sample begins in 2016. Columns 1 and 3 are PPML difference-in-differences with firm/year fixed effects and controls; columns 2 and 4 are within-firm accountant versus legal triple-differences. Panel C estimates a triple-difference comparing accountant versus legal employment within the same firm, using data stacked to two rows per firm-year with firm×year, firm×occupation, and year×occupation fixed effects. Columns 1–2 use the core accounting function (ONET 13-2011); columns 3–4 use the broader accounting and finance function. Detailed variable definitions are provided in the appendix.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 4. Hiring and Attrition Channel: Effect of Data Integration on Staff-Level Entrants and Separations

Panel A: Staff-Level Entrants

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of New Hires			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	-9.97***		-10.58***	
	(-3.80)		(-5.58)	
$2023_t \times \text{Data Integration}_{i,2022}$		-9.35***		-9.79***
		(-3.40)		(-4.93)
$2024_t \times \text{Data Integration}_{i,2022}$		-10.79***		-11.57***
		(-3.35)		(-5.08)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	-1.01	-1.02	-1.39	-1.40
	(-0.53)	(-0.54)	(-1.03)	(-1.04)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	30,841	30,841	35,022	35,022
Pseudo R-squared	0.84	0.84	0.93	0.93

Panel B: Staff-Level Separations

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of Separations			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	-1.17		-2.03	
	(-0.60)		(-1.06)	
$2023_t \times \text{Data Integration}_{i,2022}$		0.48		-1.55
		(0.24)		(-0.79)
$2024_t \times \text{Data Integration}_{i,2022}$		-3.08		-2.61
		(-1.41)		(-1.12)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	-2.18	-2.20	-2.80**	-2.81**
	(-1.48)	(-1.50)	(-2.03)	(-2.04)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	28,749	28,749	33,669	33,669
Pseudo R-squared	0.82	0.82	0.92	0.92

T-statistics in parentheses. Standard errors are clustered at the firm level. The sample spans 2010–2024. This table decomposes the staff-level employment decline documented in Table 3 into its hiring and attrition components. Panel A uses the number of staff-level new hires as the dependent variable, estimated via PPML (ppmlhdfe) with firm and year fixed effects and controls. A new hire (entrant) is defined as an employee whose first observed start date at the firm falls within the year; internal role changes within the same firm are not counted. Panel B uses the number of staff-level separations as the dependent variable (same PPML specification as Panel A). A separation is defined as an employee whose final observed end date at the firm falls within the year; employees who remain active are not counted. Columns 1 and 3 report pooled post-treatment effects; columns 2 and 4 decompose by year. Columns 1–2 use core accounting (Accountants and Auditors, O*NET 13-2011); columns 3–4 use the broader accounting and finance function. Detailed variable definitions are provided in the appendix.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 5. Within-Firm Effects of Data Integration: Manager Employment and Wages

Panel A: Manager vs. Staff Employment

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of Employees			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022} \times \text{Manager}_j$	4.537**		3.318**	
	(2.502)		(2.433)	
$2023_t \times \text{Data Integration}_{i,2022} \times \text{Manager}_j$		3.121**		2.345**
		(1.994)		(2.050)
$2024_t \times \text{Data Integration}_{i,2022} \times \text{Manager}_j$		4.242**		3.295**
		(2.330)		(2.470)
$2025_t \times \text{Data Integration}_{i,2022} \times \text{Manager}_j$		7.078***		4.821***
		(3.235)		(2.656)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022} \times \text{Manager}_j$	0.214	0.248	1.408	1.434
	(0.177)	(0.206)	(1.224)	(1.249)
Firm × Year FEs	Yes	Yes	Yes	Yes
Firm × Role FEs	Yes	Yes	Yes	Yes
Year × Role FEs	Yes	Yes	Yes	Yes
Observations	98,968	98,968	110,772	110,772
Pseudo R-squared	0.974	0.974	0.991	0.991

Panel B: Wage Effects

	Staff-Level		Manager-Level	
Dependent Variable:	Ln(Salary)			
PostGPT$_t$ × Data Integration$_{i,2022}$	0.04	0.63	7.53**	4.96**
	(0.04)	(0.63)	(2.30)	(2.05)
PostGPT$_t$ × IT Intensity$_{i,2022}$	-0.20	-0.23	0.57	0.83
	(-0.37)	(-0.54)	(0.57)	(0.79)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Metropolitan Area FEs	Yes	Yes	Yes	Yes
Seniority FEs	Yes	Yes	Yes	Yes
Task Content Controls	No	Yes	No	Yes
Controls and Post × Controls	Yes	Yes	Yes	Yes
Observations	28,041	26,196	8,256	7,855
R-squared	0.36	0.41	0.49	0.54

T-statistics in parentheses. Panel A estimates PPML triple-difference regressions of staff- vs. manager-level employment on Data Integration, using data stacked to two rows per firm-year (one manager, one staff). Manager is defined as Revelio seniority 4–6; staff as seniority 1–3. Standard errors are clustered at the firm level. The coefficient of interest captures the differential effect of Data Integration on manager versus staff-level employment within the same firm. Columns 1 and 3 report pooled post-treatment effects; columns 2 and 4 decompose by year. Columns 1–2 use core accounting (ONET 13-2011); columns 3–4 use the broader accounting and finance function. Panel B examines wage effects in job postings, where the dependent variable is the log of the midpoint posted salary (non-estimated only). Standard errors in Panel B are clustered at the metropolitan area × year level. All Panel B models include firm, year, metropolitan area, and seniority fixed effects. Columns 1 and 3 exclude task content controls; columns 2 and 4 add them. All models include firm-level controls and post × controls interactions. Detailed variable definitions are in the appendix.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 6. Robustness: Alternative Specifications

Panel A: Staff-Level Employment DiD

	Entropy Bal.		Base 2019		DI Software	
	(1) Pooled	(2) Event	(3) Pooled	(4) Event	(5) Pooled	(6) Event
Dependent Variable:	Number of Staff Employees					
$\text{PostGPT}_t \times \text{Treatment}_i$	-4.297***		-4.209***		-1.988	
	(-3.231)		(-3.374)		(-1.225)	
$2023_t \times \text{Treatment}_i$		-2.445**		-1.227		0.591
		(-2.519)		(-1.079)		(0.542)
$2024_t \times \text{Treatment}_i$		-3.837***		-2.860**		-0.994
		(-3.698)		(-2.387)		(-0.756)
$2025_t \times \text{Treatment}_i$		-4.946***		-4.151***		-2.598*
		(-4.487)		(-3.367)		(-1.683)
$\text{PostGPT}_t \times \text{IT Int.}_{i,2022}$	-0.301	-0.357	-0.429	-0.478	-1.234	-1.231
	(-0.426)	(-0.504)	(-0.425)	(-0.472)	(-1.199)	(-1.197)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
EB Weights	Yes	Yes	–	–	–	–
Observations	39,713	39,713	39,713	39,713	39,713	39,713
Pseudo R-squared	0.972	0.972	0.969	0.969	0.969	0.969

Panel B: Manager vs. Staff DDD

	(1) Baseline	(2) Base 2019	(3) DI Software	(4) Rolling Avg
Dependent Variable:	Number of Employees			
$\text{PostGPT}_t \times \text{Treatment}_i \times \text{Manager}_j$	4.564**	4.320**	4.413**	4.400**
	(2.460)	(2.362)	(2.295)	(2.197)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022} \times \text{Manager}_j$	-0.224	-0.224	0.552	-0.220
	(-0.174)	(-0.176)	(0.457)	(-0.169)
Firm × Year FEs	Yes	Yes	Yes	Yes
Firm × Role FEs	Yes	Yes	Yes	Yes
Year × Role FEs	Yes	Yes	Yes	Yes
Observations	94,080	94,080	94,080	94,080
Pseudo R-squared	0.975	0.975	0.975	0.975

T-statistics in parentheses. Standard errors are clustered at the firm level. Panel A replicates the baseline staff-level employment DiD (Table 3 Panel A) under alternative specifications, each shown in pooled and event-time form. Columns 1–2 use entropy balancing weights: firms are classified as high- or low-integration based on a within-NAICS-3 median split of Data Integration exposure, and control-group observations are reweighted to balance pre-treatment covariates, with 3–97% propensity score overlap trimming. Columns 3–4 replace the 2022-vintage treatment measure with a 2019 base year, measuring industry Data Integration exposure prior to the pandemic to rule out post-2020 confounds. Columns 5–6 replace the broad Data Integration measure with DI Software Intensity, which captures the share of employee profiles mentioning specific data integration software tools (e.g., Informatica, Talend, SSIS) rather than general integration roles. Panel B replicates the baseline manager vs. staff DDD (Table 5 Panel A col 1) under alternative treatment measures. Entropy balancing is not included in Panel B because the triple-difference specification absorbs firm×year fixed effects, which subsume the between-firm variation that entropy balancing targets; identification in the DDD comes entirely from within-firm-year differences across employee roles. Column 1 repeats the baseline (2022 Data Integration). Column 2 uses the 2019 base-year treatment. Column 3 uses DI Software Intensity. Column 4 uses a three-year rolling average of industry Data Integration exposure (2020–2022) to smooth year-to-year measurement noise. All models estimated with PPML (ppmlhdfe). Detailed variable definitions are provided in the appendix.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 7. Effect of Data Integration on Accounting Job Tasks post-GAI

Effect of Data Integration on Accounting Job Task Content (Firm-Year Means)

	Num Tasks / Total Tasks			Relative Ratios	
Dependent Variable:	Routine	Analytic	Strategic	$\frac{\text{Analytic}}{\text{Routine}}$	$\frac{\text{Strategic}}{\text{Routine}}$
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	-0.56**	0.27	0.25	0.46*	0.54**
	(-2.20)	(1.26)	(1.57)	(1.83)	(2.01)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	0.05	-0.07	-0.02	-0.04	0.08
	(0.29)	(-0.50)	(-0.15)	(-0.25)	(0.45)
Firm FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Post × Controls	Yes	Yes	Yes	Yes	Yes
Observations	10,924	10,924	10,924	10,919	10,907
R-squared	0.49	0.52	0.44	0.51	0.46

T-statistics in parentheses. Standard errors clustered at the firm level. All models aggregate job postings to firm-year means and include firm and year fixed effects, controls, and post × controls interactions. Sample restricted to core accounting job postings (O*NET 13-2011). Routine is the ratio of routine tasks to total tasks. Analytic is the ratio of analytic tasks to total tasks. Strategic is the ratio of strategic tasks to total tasks. Task definitions and mappings to each category are included in the appendix. The final two columns show the ratio of analytic to routine tasks and strategic to routine tasks, defined as the number of analytic/strategic tasks divided by the sum of routine and analytic/strategic tasks. Detailed variable definitions are in the appendix.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 8. CPA Demand and Characteristics of New Accounting Hires post-GAI

Panel A: CPA Requirement (Job Postings)

	Staff-Level		Manager-Level	
Dependent Variable:	CPA Required			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	1.18*	1.28**	-0.00	0.18
	(1.95)	(2.35)	(-0.00)	(0.19)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	-0.55	-0.59	-0.75	-0.80
	(-1.44)	(-1.61)	(-1.27)	(-1.38)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Metropolitan Area FEs	No	Yes	No	Yes
Seniority FEs	No	Yes	No	Yes
Task Content Controls	Yes	Yes	Yes	Yes
Controls and Post × Controls	Yes	Yes	Yes	Yes
Observations	231,680	231,680	74,650	74,650
R-squared	0.27	0.31	0.30	0.31

Panel B: Characteristics of New Hires

	Cert	Degrees			
Dependent Variable:	CPA	Acct	Tech	Masters	MAcc
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	0.39*	1.13***	0.24*	0.94**	0.22**
	(1.87)	(3.27)	(1.78)	(2.29)	(2.01)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	0.03	−0.46***	0.06	−0.24	−0.15***
	(0.40)	(−4.27)	(1.66)	(−1.69)	(−4.09)
Firm FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Seniority FEs	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	297,132	297,132	297,132	297,132	297,132
R-squared	0.04	0.10	0.02	0.05	0.02

T-statistics in parentheses. Panel A examines the effect of Data Integration on CPA requirements in job postings, where the dependent variable is equal to one if the posting requires a CPA, equal to 0.5 if the posting says a CPA is preferred, and 0 otherwise. Standard errors in Panel A are clustered at the metropolitan area × year level. Sample restricted to core accounting job postings (O*NET 13-2011). Columns 1 and 3 include firm and year fixed effects; columns 2 and 4 add metropolitan area and seniority fixed effects. All Panel A models control for task content (routine, analytic, and strategic task shares) and firm-level controls and post × controls interactions. Panel B uses résumé data to examine the characteristics of new accounting hires at firms in the sample. All Panel B models use OLS (reghdfe) with firm, year, and seniority fixed effects and controls; standard errors are clustered at the firm level. Column 1 examines the likelihood of a new hire having a CPA (identified from profile name and title). Columns 2–5 examine the probability of having an accounting degree, a technical degree (CS, IT, math, engineering, statistics), any master's degree, and a Master of Accounting, respectively. Detailed variable definitions are in the appendix.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Online Appendix

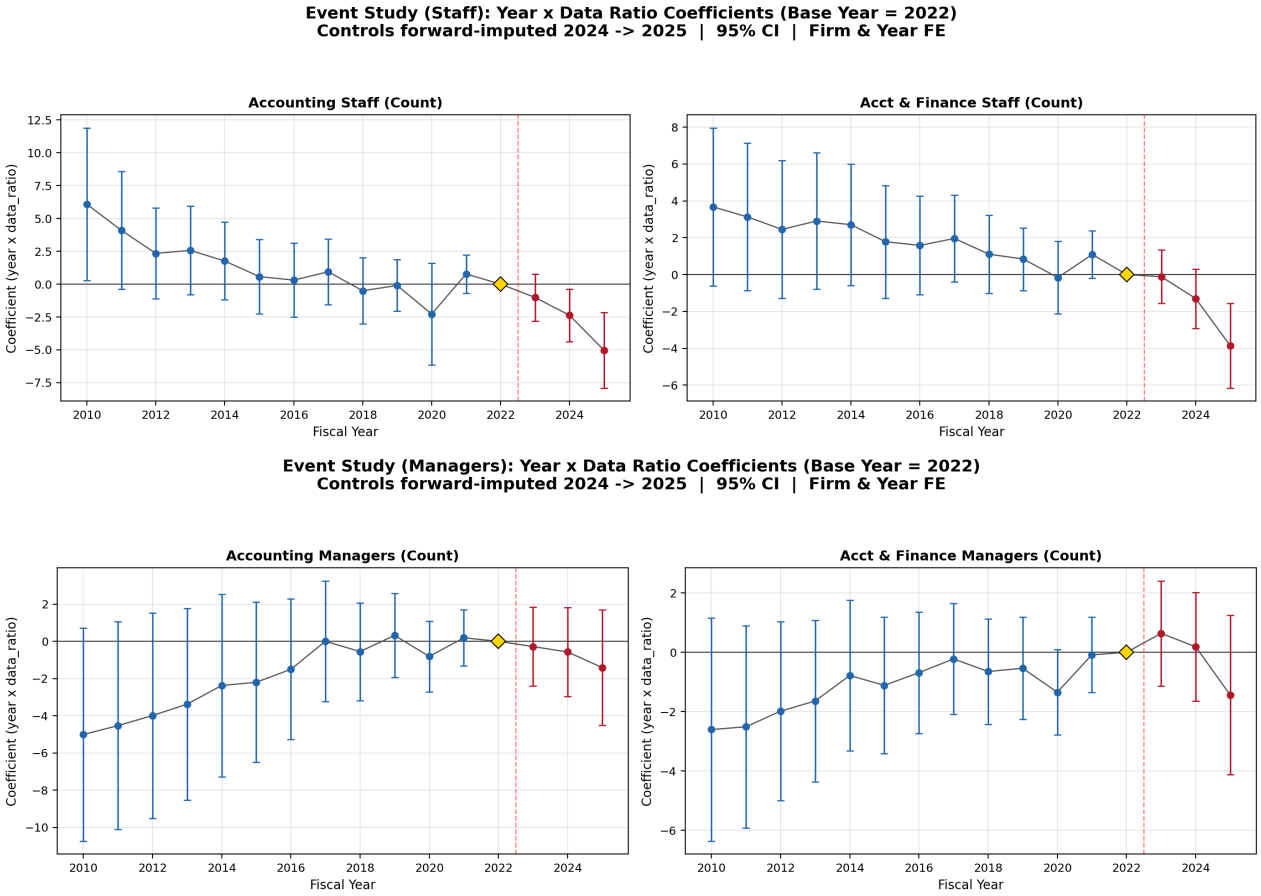


Figure OA.1. Full-window event study estimates for accounting employment. The top row plots staff-level estimates (corresponding to Figure 2 and Table 3 Panel A); the bottom row plots manager-level estimates (corresponding to Figure 3). Each row shows core accounting (left) and broader accounting and finance (right). Coefficients and 95% confidence intervals for the interaction between Data Integration and year indicators, with 2022 as the holdout year. All models estimated with PPML, firm and year fixed effects, and controls.

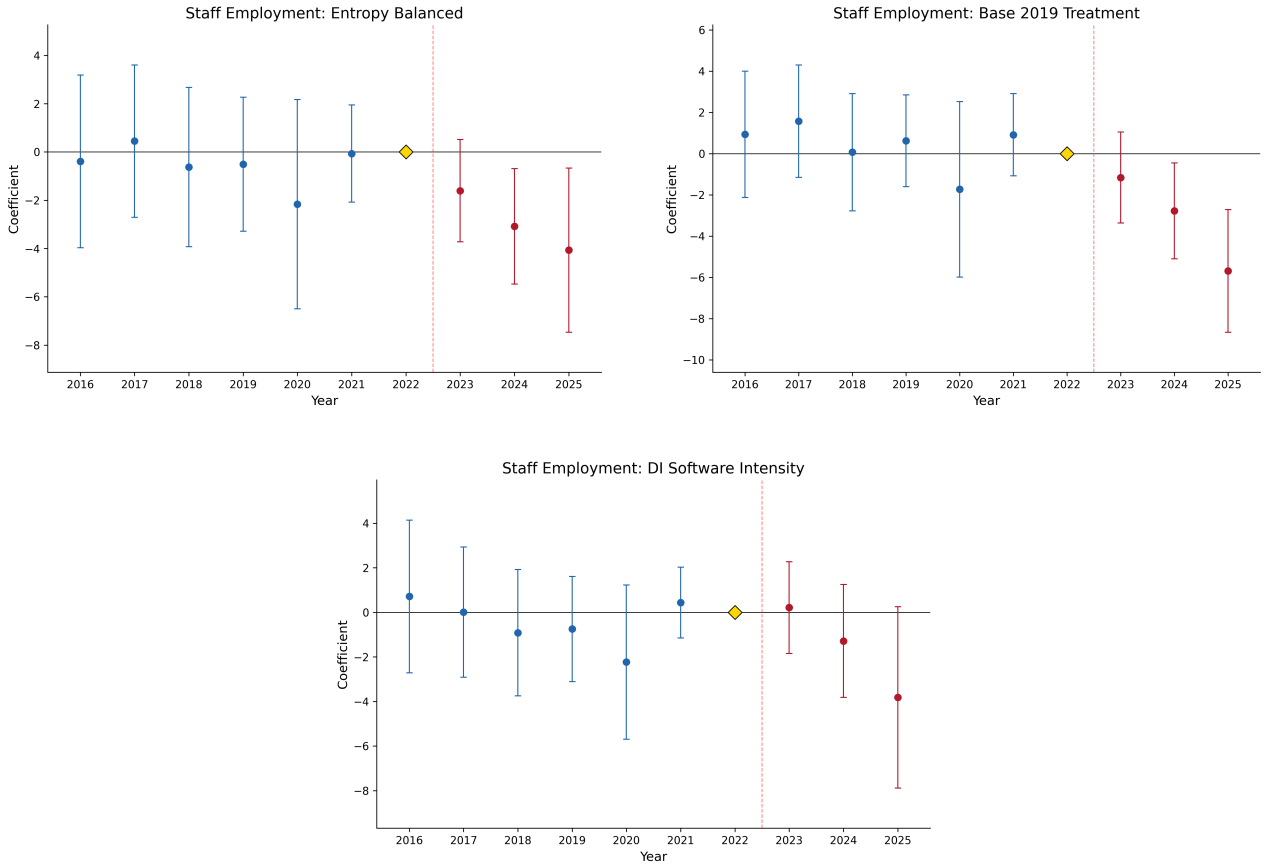


Figure OA.2. Robustness event study estimates for staff-level accounting employment under alternative specifications (corresponding to Table 6 Panel A). Top left: entropy balanced weights. Top right: 2019 base-year treatment. Bottom left: DI Software Intensity treatment. Coefficients and 95% confidence intervals with 2022 as the holdout year. The x-axis is trimmed to 2016–2025.

Table OA.1. Placebo Tests: Staff-Level Employment DiD

Panel A: 2019 Placebo (Post = 2019–2021)

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of Employees			
$\text{Post}_t \times \text{Data Integration}_{i,2019}$	-0.213		-1.384	
	(-0.183)		(-1.083)	
$2020_t \times \text{Data Integration}_{i,2019}$		-0.048		-0.668
		(-0.045)		(-0.564)
$2021_t \times \text{Data Integration}_{i,2019}$		0.258		-0.846
		(0.209)		(-0.630)
$2022_t \times \text{Data Integration}_{i,2019}$		-0.842		-2.715*
		(-0.629)		(-1.943)
$\text{Post}_t \times \text{IT Intensity}_{i,2019}$	-1.419**	-1.417**	-1.089*	-1.082*
	(-2.415)	(-2.413)	(-1.782)	(-1.773)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	34,550	34,550	36,224	36,224
Pseudo R-squared	0.968	0.968	0.985	0.985

Panel B: 2016 Placebo (Post = 2016–2018)

	Core Accounting		Broader Accounting	
Dependent Variable:	Number of Employees			
$\text{Post}_t \times \text{Data Integration}_{i,2016}$	0.000		-0.539	
	(0.000)		(-0.424)	
$2017_t \times \text{Data Integration}_{i,2016}$		1.137		0.230
		(1.074)		(0.226)
$2018_t \times \text{Data Integration}_{i,2016}$		-0.206		-0.519
		(-0.185)		(-0.419)
$2019_t \times \text{Data Integration}_{i,2016}$		-0.932		-1.323
		(-0.718)		(-0.813)
$\text{Post}_t \times \text{IT Intensity}_{i,2016}$	-2.124***	-2.120***	-1.919***	-1.917***
	(-3.911)	(-3.905)	(-3.307)	(-3.300)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	26,558	26,558	27,826	27,826
Pseudo R-squared	0.971	0.971	0.987	0.987

T-statistics in parentheses. Standard errors are clustered at the firm level. This table presents placebo tests that re-estimate the baseline staff-level employment DiD (Table 3 Panel A) at earlier treatment windows where no effect should be present. In each panel, the treatment variable (Data Integration and IT Intensity) is constructed at the indicated base year and NAICS-3 demeaned using that year's cross-section, and the sample is restricted to fiscal years up to three years after the placebo treatment date. Panel A constructs treatment at 2019 with a placebo post period of 2020–2022 (sample restricted to $\leq$ 2022). Panel B constructs treatment at 2016 with a placebo post period of 2017–2019 (sample restricted to $\leq$ 2019). Columns 1 and 3 report pooled post-treatment effects; columns 2 and 4 decompose by year. Columns 1–2 use the core accounting function (ONET 13-2011); columns 3–4 use the broader accounting and finance function. All models estimated with PPML (ppmlhdfe) with firm and year fixed effects and controls.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table OA.2. Placebo Occupations: Effect of Data Integration on Non-Accounting Employment

	Legal Staff		IT Staff		HR Staff	
	(1) Pooled	(2) Event	(3) Pooled	(4) Event	(5) Pooled	(6) Event
Dependent Variable:	Number of Employees					
$\text{PostGPT}_t \times \text{Data Integ.}_{i,2022}$	-0.016		0.926		-0.796	
	(-0.004)		(0.620)		(-0.522)	
$2023_t \times \text{Data Integ.}_{i,2022}$		2.429		-0.046		1.095
		(0.578)		(-0.030)		(0.547)
$2024_t \times \text{Data Integ.}_{i,2022}$		0.357		1.683		-0.940
		(0.078)		(1.046)		(-0.570)
$2025_t \times \text{Data Integ.}_{i,2022}$		-4.130		1.503		-4.530*
		(-0.864)		(0.768)		(-1.740)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	-3.957**	-4.076**	-3.549***	-3.532***	-1.458*	-1.536*
	(-2.171)	(-2.235)	(-6.044)	(-6.023)	(-1.757)	(-1.819)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,333	7,333	33,395	33,395	37,130	37,130
Pseudo R-squared	0.445	0.445	0.972	0.972	0.955	0.955

T-statistics in parentheses. Standard errors are clustered at the firm level. This table re-estimates the baseline staff-level employment DiD (Table 3 Panel A) using placebo dependent variables that should not be directly affected by Data Integration exposure. Columns 1–2 use legal staff employee counts as the dependent variable. Columns 3–4 use non-data IT staff employee counts. Columns 5–6 use core HR staff employee counts. The treatment variables, fixed effects, and controls are identical to Table 3 Panel A. If the baseline results reflect a causal effect of Data Integration on accounting employment specifically, we should not observe significant effects on these placebo occupations. All models estimated with PPML (ppmlhdfe).

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table OA.3. Effect of Data Integration on Data Employee Employment

	Data Staff		Data Managers	
	(1) Pooled	(2) Event	(3) Pooled	(4) Event
Dependent Variable:	Number of Data Employees			
$\text{PostGPT}_t \times \text{Data Integration}_{i,2022}$	0.894		-0.731	
	(0.404)		(-0.369)	
$2023_t \times \text{Data Integration}_{i,2022}$		1.324		0.074
		(0.750)		(0.042)
$2024_t \times \text{Data Integration}_{i,2022}$		1.250		-0.705
		(0.588)		(-0.360)
$2025_t \times \text{Data Integration}_{i,2022}$		-0.443		-2.297
		(-0.130)		(-0.760)
$\text{PostGPT}_t \times \text{IT Intensity}_{i,2022}$	-3.432***	-3.451***	-6.733***	-6.764***
	(-3.079)	(-3.102)	(-5.327)	(-5.353)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	30,022	30,022	25,288	25,288
Pseudo R-squared	0.964	0.964	0.953	0.953

T-statistics in parentheses. Standard errors are clustered at the firm level. This table estimates the effect of industry Data Integration exposure on the employment of data integration employees themselves. Columns 1–2 use staff-level (non-managerial) data employee counts as the dependent variable. Columns 3–4 use managerial data employee counts. The treatment variables, fixed effects, and controls are identical to Table 3 Panel A. All models estimated with PPML (ppmlhdfe).

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table OA.4. Additional Robustness: Alternative Fixed Effects, Functional Form, and IT Placebo

	(1) ERP × Year	(2) Ind. × Year	(3) Asinh OLS	(4) IT Only
Dependent Variable:	Number of Staff Employees			
PostGPT$_t$ × Data Integration$_{i,2022}$	-3.328***	-3.340***	-1.796***	
	(-2.620)	(-2.984)	(-2.898)	
PostGPT$_t$ × IT Intensity$_{i,2022}$	-1.097	-0.760	-0.760*	-1.646
	(-1.075)	(-1.002)	(-1.796)	(-1.562)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	–	Yes	Yes
ERP × Year FEs	Yes	–	–	–
Industry × Year FEs	–	Yes	–	–
Controls	Yes	Yes	Yes	Yes
Estimator	PPML	PPML	OLS	PPML
Observations	40,846	40,740	48,022	40,846
(Pseudo/Within) R-squared	0.969	0.970	0.200	0.968

T-statistics in parentheses. Standard errors are clustered at the firm level. The dependent variable is the number of staff-level accounting employees (columns 1–2 and 4 estimated with PPML; column 3 uses the inverse hyperbolic sine transformation estimated with OLS). This table presents additional robustness checks for the baseline staff-level accounting employment DiD (Table 3 Panel A col 1). Column 1 adds ERP adopter × year fixed effects to control for differential employment trends among firms whose employees report enterprise resource planning software (e.g., SAP, Oracle, Microsoft Dynamics) on their profiles, ensuring that the Data Integration effect is not confounded by broader ERP-driven workforce changes. Column 2 replaces year fixed effects with NAICS-3 industry × year fixed effects to absorb industry-specific time trends. Column 3 uses the inverse hyperbolic sine of staff employee counts as the dependent variable and estimates via OLS (reghdfe) to verify that the results are not driven by the PPML functional form assumption. Column 4 includes only IT Intensity (omitting Data Integration) to confirm that the baseline effect is specific to data integration exposure rather than general IT orientation. All specifications include firm fixed effects and controls.
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table OA.5. Full Coefficient Estimates: Staff-Level Employment DiD

	Core Accounting		Broader Accounting	
	(1) Pooled	(2) Event	(3) Pooled	(4) Event
Dependent Variable:	Number of Staff Employees			
Treatment Variables				
PostGPT$_t$ × Data Integr.$_{i,2022}$	-3.317***		-2.917***	
	(-2.611)		(-2.870)	
2023_t × Data Integr.$_{i,2022}$		-1.849		-1.572
		(-1.595)		(-1.552)
2024_t × Data Integr.$_{i,2022}$		-3.224**		-2.781***
		(-2.553)		(-2.639)
2025_t × Data Integr.$_{i,2022}$		-5.871***		-5.364***
		(-3.527)		(-4.390)
PostGPT$_t$ × IT Intensity$_{i,2022}$	-1.054	-1.087	-1.493*	-1.536**
	(-1.019)	(-1.049)	(-1.922)	(-1.967)
Level Controls				
Firm Age	0.032***	0.032***	0.037***	0.037***
	(5.008)	(5.031)	(6.017)	(6.031)
Log(Total Assets)	0.174***	0.175***	0.191***	0.192***
	(7.870)	(7.914)	(9.147)	(9.200)
CapEx / Assets	0.369***	0.369***	0.432***	0.432***
	(2.749)	(2.745)	(3.029)	(3.022)
Return on Assets	0.137**	0.136**	0.103**	0.102**
	(2.202)	(2.200)	(2.189)	(2.190)
Sales Growth	-0.000	-0.000	-0.000	-0.000
	(-0.584)	(-0.583)	(-1.291)	(-1.294)
Leverage	0.108	0.108	0.006	0.006
	(1.342)	(1.344)	(0.093)	(0.096)
Book-to-Market	-0.008	-0.008	-0.006	-0.006
	(-1.256)	(-1.267)	(-0.961)	(-0.976)
Average Stock Returns	-0.596***	-0.594***	-0.492***	-0.491***
	(-8.684)	(-8.678)	(-6.321)	(-6.325)
Accounting Emp. Share	9.598***	9.599***	6.656***	6.653***
	(9.402)	(9.383)	(8.087)	(8.080)

Full Coefficient Estimates (continued)

	Core Accounting		Broader Accounting	
	(1) Pooled	(2) Event	(3) Pooled	(4) Event
Post-Period × Controls				
Post × Log(Assets)	0.007	0.007	-0.011***	-0.011***
	(1.216)	(1.256)	(-2.620)	(-2.603)
Post × Leverage	-0.113**	-0.114**	-0.104***	-0.104***
	(-2.094)	(-2.138)	(-2.794)	(-2.807)
Post × Firm Age	-0.002	-0.002	-0.002	-0.002
	(-0.963)	(-0.964)	(-0.705)	(-0.705)
Post × CapEx/Assets	0.004	0.001	0.213	0.214
	(0.013)	(0.003)	(0.736)	(0.735)
Post × ROA	0.094	0.083	0.080	0.070
	(1.165)	(1.051)	(1.056)	(0.947)
Post × Book-to-Market	-0.004	-0.004	-0.002	-0.002
	(-1.136)	(-1.173)	(-0.535)	(-0.562)
Post × Avg. Returns	0.765***	0.798***	0.715***	0.735***
	(3.073)	(3.158)	(3.042)	(3.080)
Post × Sales Growth	0.000	0.000	0.001*	0.001*
	(1.277)	(1.230)	(1.808)	(1.763)
Post × Acct. Emp. Share	0.479***	0.478***	0.595***	0.591***
	(6.156)	(6.169)	(7.800)	(7.691)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	40,846	40,846	42,862	42,862
Pseudo R-squared	0.969	0.969	0.985	0.985

T-statistics in parentheses. Standard errors are clustered at the firm level. This table reports the full set of coefficient estimates from the staff-level employment DiD regressions presented in Table 3 Panel A. The specifications are identical; the only difference is that all control variable coefficients are shown. All models estimated with PPML (ppmlhdfe) with firm and year fixed effects.

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

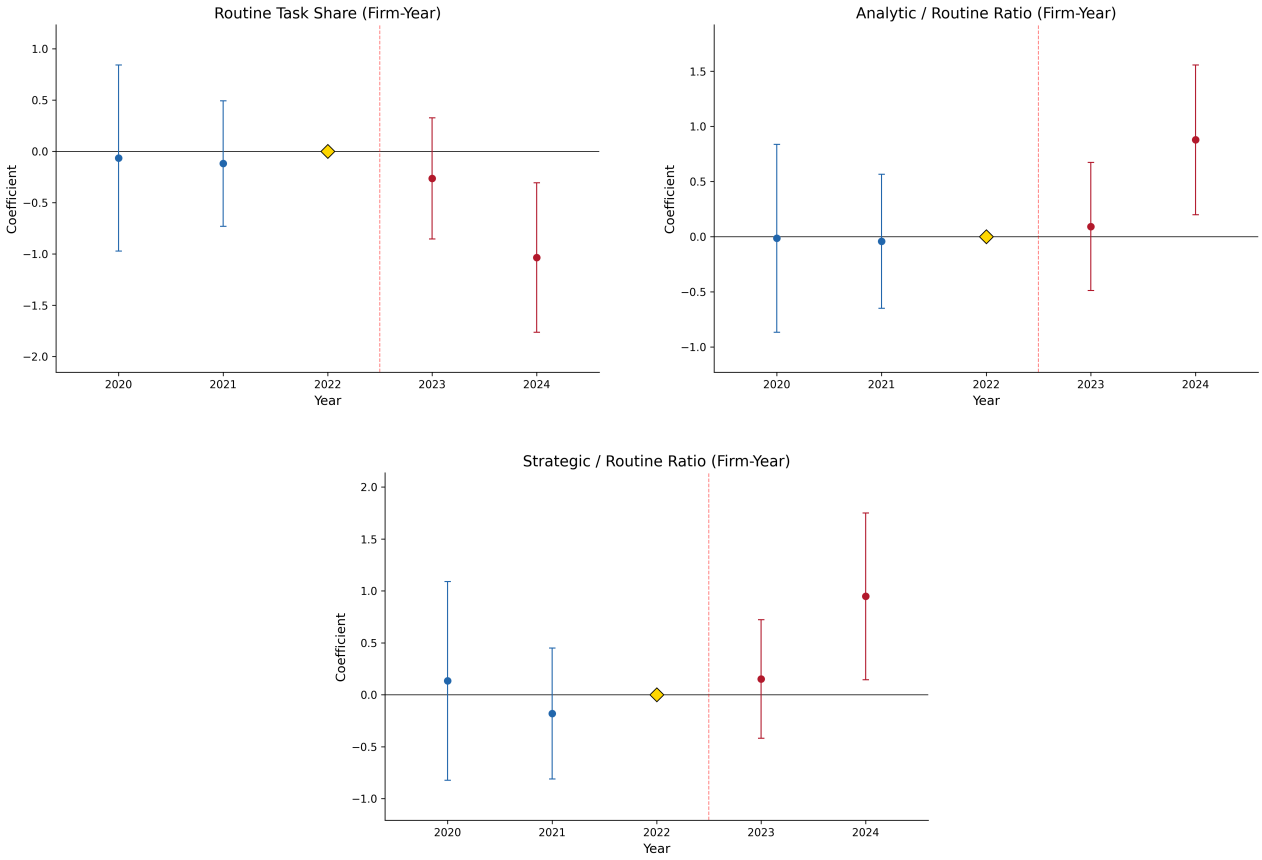


Figure OA.3. Event study estimates for firm-year task content regressions (corresponding to Table 7). Top left: routine task share (Bucket 1 / Total). Top right: analytic-to-routine ratio (Bucket 3 / (Bucket 3 + Bucket 1)). Bottom left: strategic-to-routine ratio (Bucket 4 / (Bucket 4 + Bucket 1)). Each point represents the coefficient on the interaction of the year indicator with the firm's 2022 Data Integration measure. The holdout year is 2022 (diamond). Blue markers indicate pre-treatment years; red markers indicate post-treatment years. 95% confidence intervals shown, with standard errors clustered at the firm level.

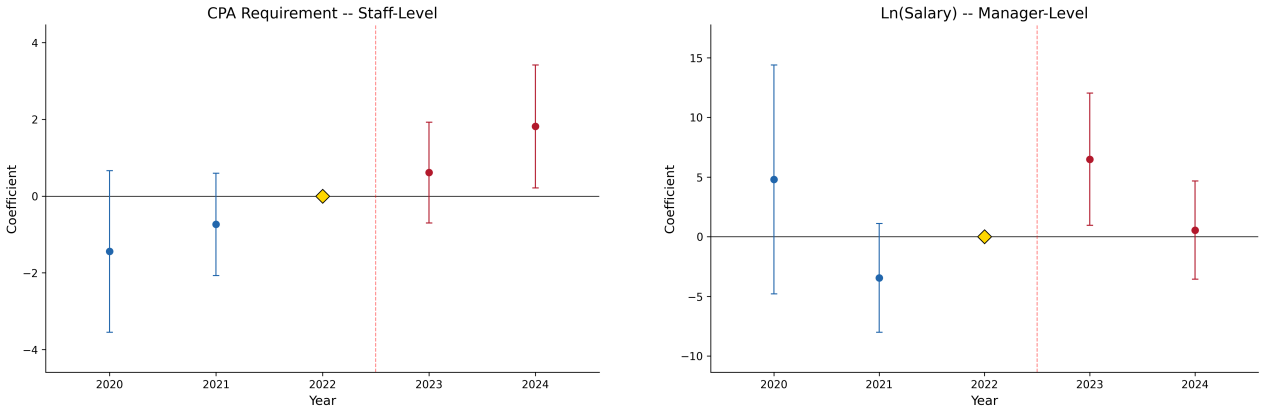


Figure OA.4. Event study estimates for CPA demand and wage regressions (corresponding to Table 8). Left: CPA requirement (binary) for staff-level accountant postings, with firm, year, metropolitan area, and seniority fixed effects, controlling for task content. Right: log posted salary for manager-level accountant postings (disclosed salaries only), with full fixed effects and task content controls. Each point represents the coefficient on the interaction of the year indicator with the firm's 2022 Data Integration measure. The holdout year is 2022 (diamond). Blue markers indicate pre-treatment years; red markers indicate post-treatment years. 95% confidence intervals shown, with standard errors clustered at the metropolitan area $\times$ year level.

OA.1: Detailed Occupational Titles

To map to O*NET SOC codes, Revelio uses a proprietary classification system based on a text embeddings model. They establish a list of 1500 job titles, and then provide a mapping between each of these job titles and the SOC codes. In this section, I list out the specific job titles (variable name: role_k1500 in Revelio's individual positions database) associated with each of my O*NET codes used in my analysis (for accounting, finance, and data management roles).

15-1242: Database Administrators

Integration, Business Intelligence Developer, Business System Analyst, Business Systems, Business Systems Analyst, Business Systems Consultant, Cloud Architect, Cloud Engineer, Database Administrator, Oracle Database Administrator, SAP Basis Administrator, SAP FICO Consultant, SQL Server DBA

15-1243: Database Architects

Data Architect, Data Engineer, Database, Database Analyst, Database Developer, Database Engineer, DBA, ETL Developer, Integration Engineer, Oracle DBA, Oracle Developer

13-2011: Accountants and Auditors

Tax Consultant, Accountant, Accounting, Accounting Analyst, Accounting Finance, Audit, Auditor, Cost Accountant, Cost Accounting, Finance Accounting, Financial Accountant, Financial Accounting, Financial Reporting, Financial Reporting Analyst, Fraud, Fund Accountant, Fund Accounting, GL Accountant, Ledger Accountant, Project Accountant, Revenue Accountant, Tax, Tax Accountant, Tax Preparer

13-2051: Financial and Investment Analysts

Finance Accounts, Account Analyst, Account Officer, Accounts Officer, Business Finance, Commercial Analyst, Commercial Finance, Control Analyst, Cost Analyst, Finance Analyst,

Finance Business, Financial, Financial Analysis, Financial Analyst, Financial Analyst FP&A, Financial Planning, Financial Planning Analysis, Financial Planning Analyst, FP&A, Pricing Analyst, Project Finance, Revenue, Revenue Analyst, Tax Analyst, Treasury, Treasury Analyst

13-2054: Financial Risk Specialists

Credit Risk, Credit Risk Analyst, Market Risk, Risk, Risk Analyst, Risk Consultant, Risk Officer

15-2011: Actuaries

Actuarial, Actuarial Analyst, Actuary, Quantitative Analyst

11-3031: Financial Managers

Business Controller, Capital Markets, Control, Controller, Controlling, Cost Controller, Equity, Equity Analyst, Finance, Finance Administration, Finance Controller, Finance Officer, Finance Operations, Financial Controller, Financial Officer, Fund Administrator, Investment, Investment Analyst, Investment Banking, Investment Banking Analyst, Investments, Investor Relations, Leadership Development, M&A, Markets, Mergers Acquisitions, Plant Controller, Portfolio, Portfolio Analyst, Project Controller, Trading

43-3021: Billing and Posting Clerks

A/R, Billing, Billing Analyst, Reimbursement, Transaction Processing

43-3031: Bookkeeping, Accounting, and Auditing Clerks

Accounts Payable, Accounts Payable Analyst, Accounts Receivable, Bookkeeper

OA.2: Classification of Accounting Tasks into Routine, Analytic, Strategic, and Other

Classification of O*NET Accounting Tasks

Task Description	**Category**
Routine Tasks	
Prepare adjusting journal entries.	Routine
Review accounts for discrepancies and reconcile differences.	Routine
Process invoices for payment.	Routine
Establish tables of accounts and assign entries to proper accounts.	Routine
Inspect cash on hand, notes receivable and payable, negotiable securities, and canceled checks to confirm records are accurate.	Routine
Analytical Tasks	
Analyze business operations, trends, costs, revenues, financial commitments, and obligations to project future revenues and expenses or to provide advice.	Analytic
Examine and evaluate financial and information systems, recommending controls to ensure system reliability and data integrity.	Analytic
Conduct pre-implementation audits to determine if systems and programs under development will work as planned.	Analytic
Collect and analyze data to detect deficient controls, duplicated effort, extravagance, fraud, or non-compliance with laws, regulations, and management policies.	Analytic
Prepare, analyze, or verify annual reports, financial statements, and other records, using accepted accounting and statistical procedures to assess financial condition and facilitate financial planning.	Analytic
Develop, maintain, or analyze budgets, preparing periodic reports that compare budgeted costs to actual costs.	Analytic
Develop, implement, modify, and document recordkeeping and accounting systems, making use of current computer technology.	Analytic
Advise clients in areas such as compensation, employee health care benefits, the design of accounting or data processing systems, or long-range tax or estate plans.	Analytic

Classification of O*NET Accounting Tasks (Continued)

Task Description	Category
Strategic Tasks	
Prepare detailed reports on audit findings.	Strategic
Report to management about asset utilization and audit results, and recommend changes in operations and financial activities.	Strategic
Supervise auditing of establishments, and determine scope of investigation required.	Strategic
Confer with company officials about financial and regulatory matters.	Strategic
Report to management regarding the finances of establishment.	Strategic
Examine whether the organization's objectives are reflected in its management activities, and whether employees understand the objectives.	Strategic
Direct activities of personnel engaged in filing, recording, compiling, and transmitting financial records.	Strategic
Other Tasks	
Inspect account books and accounting systems for efficiency, effectiveness, and use of accepted accounting procedures to record transactions.	Other
Examine records and interview workers to ensure recording of transactions and compliance with laws and regulations.	Other
Audit payroll and personnel records to determine unemployment insurance premiums, workers' compensation coverage, liabilities, and compliance with tax laws.	Other
Review taxpayer accounts, and conduct audits on-site, by correspondence, or by summoning taxpayer to office.	Other
Compute taxes owed and prepare tax returns, ensuring compliance with payment, reporting, or other tax requirements.	Other
Prepare, examine, or analyze accounting records, financial statements, or other financial reports to assess accuracy, completeness, and conformance to reporting and procedural standards.	Other
Examine inventory to verify journal and ledger entries.	Other
Evaluate taxpayer finances to determine tax liability, using knowledge of interest and discount rates, annuities, valuation of stocks and bonds, and amortization valuation of depletable assets.	Other
Review data about material assets, net worth, liabilities, capital stock, surplus, income, or expenditures.	Other

This table provides the mapping between the 29 unique O*NET tasks for accountants and auditors and their categorization into routine, analytic, strategic, and other. Other tasks are the remaining tasks not easily placed into any of the three categories. These categories were created using a large language model.

OA.3: Generative AI Keyword Dictionary

For my validation analysis, I create a keyword of words related specifically to *generative* AI. The goal of the keyword list is to accurately pick up keywords that are specific to generative AI and not broader trends in technology and other forms of artificial intelligence. I used domain knowledge and the assistance of 3 generative AI tools (Claude, Gemini, and ChatGPT) to create this list. Minimizing false positives was one of the objectives of this keyword list, but that comes at the likely tradeoffs of type 2 errors. The list is case-insensitive.

Keyword List:

"generative ai", "generative artificial intelligence", "genai", "gen ai", "large language model", "large language models", "foundation model", "multimodal model", "foundation models", "multimodal models", "retrieval-augmented generation", "retrieval augmented generation", "rag pipeline", "rag architecture", "chatgpt", "chat gpt", "openai", "azure openai", "azure openai service", "gpt-3.5", "gpt3.5", "gpt-4", "gpt4", "gpt-4o", "gpt4o", "gpt-4.1", "gpt4.1", "gpt-4.5", "gpt4.5", "gpt-5", "gpt5", "google bard", "bard ai", "google gemini", "gemini ai", "gemini 1.5", "gemini 2.5", "palm 2", "palm2", "anthropic claude", "claude 2", "claude 3", "claude 3.7", "claude 3.5", "llama 2", "llama2", "llama 3", "llama3", "mistral-7b", "mistral", "mixtral", "mixtral 8x7b", "grok 3", "grok-3", "qwen 3", "deepseek", "deep seek" "microsoft copilot", "m365 copilot", "github copilot", "dall-e", "dalle", "dall-e 2", "stable diffusion", "sdxl", "midjourney", "sora", "langchain", "LLM", "LLMs", "generative pre-trained transformer", "generative pretrained transformer", "prompt engineering", "text-to-image", "Gemini Pro", "Gemini Ultra", "Gemini Flash", "Claude Opus", "Claude Sonnet", "claude haiku", "AI agent", "agentic AI", "generative chatbot", "transformer model", "transformer models", "transformer architecture", "diffusion model", "diffusion models", "zero-shot", "few-shot", "zero shot", "few shot", "instruction tuning", "AI alignment", "vector database", "vector databases", "vector embedding", "vector embeddings", "token limit", "context window", "Hugging Face", "huggingface", "OpenAI API", "Pinecone", "artificial general intelligence"